

**Learners' perceptions of language proficiency, language test-taking
strategies and emotional regulation in a test-taking context:**

A case study in an Egyptian EFL context

DEENA BORAIE

**A thesis submitted in partial fulfilment of the
requirements of the University of
Wolverhampton for the degree of Doctor of
Philosophy**

July 2003

This work or any part thereof has not previously been presented in any form to the University or to any other body whether for the purposes of assessment, publication or for any other purpose (unless previously indicated). Save for any express acknowledgements, references and/or bibliographies cited in the work, I confirm that the intellectual content of the work is the result of my own efforts and of no other person.

The right of Deena Boraie to be identified as author of this work is asserted in accordance with ss. 77 and 78 of the Copyright, Designs and Patents Act 1988. At this date copyright is owned by the author.

Signed 

Date 

ACKNOWLEDGEMENTS

In carrying out this research study, innumerable colleagues, friends and members of my family have supported me throughout the whole process. Without their support, this thesis would never have been completed. I have so many people to thank that I cannot do justice to them all.

First and foremost, my deepest gratitude goes to my supervisor, Dr. Jo Allan, Associate Dean of the School of Education at the University of Wolverhampton. She has helped me become a far better researcher than I ever was and she has taught me the significance of examining research design at a much deeper level. Although we were miles apart, she gave freely of her time with guidance, support and encouragement. I feel privileged to have had her as my supervisor in this dissertation.

My heartfelt thanks go to my mother, Rhoda Hussein and my father, Aly Boraie for their love and support and who never failed to be there for me. I hope I have made them proud of me.

I am indebted to three friends and colleagues: Omneya Kassabgy for her support and for her insightful comments throughout the thesis; to Fred Perry and Russanne Hozayin for their time in reviewing my work providing me with invaluable feedback.

I also wish to express my appreciation to another three wonderful colleagues. I want to thank Nagwa Kassabgy for her time and patience in accepting to do the painstaking task of going through the think aloud data and coding them. I am grateful to Ahmed Dewidar for his help in the back translation of the questionnaire I used in the study. I am deeply grateful to Ibrahim Said for his assistance and support in enabling me to collect the data.

I wish to express my gratitude to all my friends and colleagues at the Center for Adult & Continuing Education (CACE) of the American University in Cairo for their support and encouragement. I would like to thank Robert Younghouse and Magda Laurence and to express my sincere thanks to those who helped with the data collection: Bassem Boulos, Wissa Akhnoukh and Mary Tawadros.

Finally, my special thanks go to the wonderful learners at CACE who gave of their time to spend two hours thinking aloud while they took a language test and who stayed after a challenging test to complete a questionnaire.

ABSTRACT

This study investigated adult EFL learners' perceptions of English language proficiency, identified their test-taking strategies and emotional regulation processes during test-taking and explored the relationship between test-takers' reported use of test-taking strategies and emotional regulation on the one hand and their performance on an English language placement test on the other. The study was conducted in the Center for Adult & Continuing Education at the American University in Cairo.

The rationale for this study was two fold. Firstly, given that there is a variety of interpretations of the construct of English language proficiency, no empirically derived definition of language proficiency was available for this particular context. Secondly, a lack of research was found investigating test-taking strategies, emotional regulation and test performance within a process approach. Expanding on the Bachman & Palmer (1996) model, a Language Testing Processing (LTP) model was proposed.

A mixed methods approach was used integrating qualitative and quantitative methods in different stages of the study. The context-specific definition of the construct of language proficiency was based on data collected from 36 learners using a semi-structured interview and from 41 teachers using an open-ended questionnaire. Test-taking strategies and emotional regulation processes were identified from think aloud data obtained from 12 test-takers who were asked to verbalize their thoughts and feelings while they took a placement test. Based on the think aloud data and the literature, the Test-Taking Strategies Questionnaire (TTSQ) was designed to investigate the relationship between test-taking strategies, emotional regulation and language test performance. The TTSQ was administered to 497 test-takers after they completed the placement test.

Correlations, analysis of variance and discriminant analysis showed that emotional regulation influences the selection of particular test-taking strategies, which in turn is associated with performance on a language placement test. The LTP model was supported and further refined by the think aloud and quantitative data.

The insights gained on perceptions of language proficiency, test-taking strategies and affective factors that influence test performance are discussed and the implications of these results on curriculum designers, test developers and teachers are presented.

DEDICATION
TO BASIL AND NEVINE

TABLE OF CONTENTS

	<u>PAGE</u>
ACKNOWLEDGEMENTS	
ABSTRACT	
LIST OF TABLES	xii
LIST OF FIGURES	xvii
LIST OF APPENDICES	xviii
CHAPTER 1: INTRODUCTION	1
1.1 BACKGROUND OF THE STUDY	1
1.2 PURPOSE OF THE STUDY	3
1.3 RESEARCH QUESTIONS	3
1.3.1 Research Question 1	3
1.3.2 Research Question 2	3
1.3.3 Research Question 3	4
1.4 OVERVIEW OF THE THESIS	4
CHAPTER 2: LITERATURE REVIEW	6
2.1 VALIDITY & LANGUAGE TESTING MODELS	6
2.1.1 Importance of Validity	7
2.1.2 Changing Concepts of Validity	7
2.1.3 Unified Validity Concept	8
2.1.4. Validation Research in Language Testing	9
2.1.5 Models in Language Testing	12
2.1.6 Scales of Proficiency	13
2.1.7 Models of Language Proficiency	15
2.1.7.1 Unitary Competence Hypothesis	15
2.1.7.2 Munby’s Model	16
2.1.7.3 Canale and Swain Model	17

	<u>PAGE</u>
2.1.7.4 Bachman Model	18
2.1.7.5 Bachman and Palmer's Model	21
2.1.7.6 Summary of Models of Language Proficiency	26
2.1.8 Research Question 1	28
2.2 RESEARCH IN TEST-TAKING STRATEGIES	29
2.2.1 Defining Test-taking Strategies	29
2.2.2 Identifying Test-taking Strategies	30
2.2.3 Studies on Test-taking Strategies	31
2.2.3.1 Reading Comprehension Tests	32
2.2.3.2 Listening Tests	41
2.2.3.3 Proficiency Tests	41
2.2.4 Summary of Test-taking Strategy Studies Reviewed	43
2.3 AFFECTIVE FACTORS AND TEST PERFORMANCE	47
2.4 RESEARCH IN LANGUAGE SELF-ASSESSMENT	49
2.4.1 Validation Studies Supporting Self-Assessment	50
2.4.2 Validation Studies not Supporting Self-Assessment	50
2.4.3 Accuracy of Self-Assessment	52
2.4.4 Design of Self-Assessment Instruments and Benchmarking	53
2.4.5 Process of Self-Assessment	53
2.5 SELF-REGULATION WHILE TAKING TESTS	55
2.5.1 Definition of Terms in Self-Regulation	55
2.5.2 Self-Regulation Models	57
2.5.3 Emotional Regulation in Test Situations	63
2.5.4 Research Questions 2 & 3	71
2.6 PROPOSED LANGUAGE TESTING PROCESSING (LTP) MODEL	68
2.7 DEFINITION OF TERMS	71
2.8 SUMMARY	71

	<u>PAGE</u>
3.6.2 Data Collection Procedures for the Open-Ended Questionnaires	102
3.6.3 Data Collection Procedures for the Think Aloud Verbal Protocols	102
3.6.3.1 Description of the Test Task	104
3.6.4 Data Collection Procedures for Administering the TTSQ	105
3.7 GENERALIZABILITY AND RELATABILITY	106
3.8 VALIDITY	108
3.9 RELIABILITY	112
3.9.1 Procedural Reliability	112
3.9.1.1 Reliability of the Think Aloud Verbal Reports	112
3.9.2 Reliability of the TTSQ	113
3.9.3 Inter-Coder Reliability for Coding the Qualitative Data	113
3.10 ETHICS	114
3.11 EVALUATING THE RESEARCH DESIGN	115
3.12 ANALYSIS OF DATA	118
3.12.1 Analyzing Quantitative Data	118
3.12.1.1 Factor Analysis	118
3.12.1.2 Analysis of Variance (ANOVA)	118
3.12.1.3 Correlations	119
3.12.1.4 Discriminant Analysis	119
3.12.2 Analyzing Qualitative Data	120
3.12.2.1 Analyzing Research Question 1 Qualitative Data	121
3.12.2.2 Analyzing Research Question 2 Qualitative Data	122
3.13 SUMMARY	128

	<u>PAGE</u>
CHAPTER 4: RESULTS & FINDINGS FOR RESEARCH QUESTION 1	129
4.1 RESEARCH QUESTION 1	129
4.2 FINDINGS ON MOTIVES FOR STUDYING ENGLISH AND DIFFERENT CONTEXTS OF LANGUAGE USE	129
4.3 FINDINGS ON DEFINING THE CONSTRUCT OF LANGUAGE PROFICIENCY	134
4.4 FINDINGS ON THE RELATIONSHIP BETWEEN THE CONSTRUCT OF LANGUAGE PROFICIENCY AND THE CONTEXTS OF LANGUAGE USE	138
4.5 DEFINITION OF THE CONSTRUCT OF ENGLISH LANGUAGE PROFICIENCY OF ADULT EFL LEARNERS AT CACE, AUC	139
4.6 SUMMARY OF FINDINGS AND OUTCOMES ON RESEARCH QUESTION 1	141
4.6.1 Research Question 1	141
4.6.2 Research Question 1.1	141
4.6.3 Research Question 1.2	142
4.6.4 Research Question 1.3	142
CHAPTER 5: RESULTS & FINDINGS FOR RESEARCH QUESTION 2	143
5.1 RESEARCH QUESTION 2	143
5.2 FINDINGS FROM THE THINK ALOUD DATA: IDENTIFYING TEST-TAKING STRATEGIES AND FIT OF THE LTP MODEL	143

		<u>PAGE</u>
5.2.1	Test-taker A	144
5.2.1.1	Writing	144
5.2.1.2	Listening	144
5.2.1.3	Grammar	145
5.2.1.4	Reading	145
5.2.1.5	Affective Responses	146
5.2.1.6	Fit of the LTP Model	146
5.2.2	Test-taker B	148
5.2.2.1	Writing	148
5.2.2.2	Listening	148
5.2.2.3	Grammar	149
5.2.2.4	Reading	149
5.2.2.5	Affective Responses	150
5.2.2.6	Fit of the LTP Model	150
5.2.3	Test-taker C	152
5.2.3.1	Writing	152
5.2.3.2	Listening	152
5.2.3.3	Grammar	153
5.2.3.4	Reading	153
5.2.3.5	Affective Responses	154
5.2.3.6	Fit of the LTP Model	154
5.2.4	Test-taker D	155
5.2.4.1	Writing	155
5.2.4.2	Listening	156
5.2.4.3	Grammar	156
5.2.4.4	Reading	157
5.2.4.5	Affective Responses	157
5.2.4.6	Fit of the LTP Model	157
5.2.5	Test-taker E	158
5.2.5.1	Writing	158
5.2.5.2	Listening	159
5.2.5.3	Grammar	159
5.2.5.4	Reading	160
5.2.5.5	Affective Responses	160
5.2.5.6	Fit of the LTP Model	160

	<u>PAGE</u>
5.2.6 Test-taker F	162
5.2.6.1 Writing	162
5.2.6.2 Listening	162
5.2.6.3 Grammar	163
5.2.6.4 Reading	163
5.2.6.5 Affective Responses	163
5.2.6.6 Fit of the LTP Model	164
5.2.7 Test-taker G	165
5.2.7.1 Writing	165
5.2.7.2 Listening	166
5.2.7.3 Grammar	166
5.2.7.4 Reading	167
5.2.7.5 Affective Responses	167
5.2.1.6 Fit of the LTP Model	168
5.2.8 Test-taker H	168
5.2.8.1 Writing	168
5.2.8.2 Listening	169
5.2.8.3 Grammar	169
5.2.8.4 Reading	170
5.2.8.5 Affective Responses	170
5.2.8.6 Fit of the LTP Model	171
5.2.9 Test-taker I	171
5.2.9.1 Writing	171
5.2.9.2 Listening	171
5.2.9.3 Grammar	172
5.2.9.4 Reading	172
5.2.9.5 Affective Responses	172
5.2.9.6 Fit of the LTP Model	173
5.2.10 Test-taker J	173
5.2.10.1 Writing	173
5.2.10.2 Listening	173
5.2.10.3 Grammar	174
5.2.10.4 Reading	174
5.2.10.5 Affective Responses	174
5.2.10.6 Fit of the LTP Model	175

	<u>PAGE</u>
5.2.11 Test-taker K	175
5.2.11.1 Writing	175
5.2.11.2 Listening	176
5.2.11.3 Grammar	176
5.2.11.4 Reading	176
5.2.11.5 Affective Responses	176
5.2.11.6 Fit of the LTP Model	177
5.2.12 Test-taker L	177
5.2.12.1 Writing	177
5.2.12.2 Listening	178
5.2.12.3 Grammar	178
5.2.12.4 Reading	178
5.2.12.5 Affective Responses	178
5.2.12.6 Fit of the LTP Model	179
5.3 SUMMARY OF FINDINGS FROM THE THINK ALOUD DATA	179
5.3.1 Test-taking Strategies	179
5.3.2 Emotions and Emotional Regulation Processes	182
5.3.2.1 Assessment: Metacognitive Strategies	182
5.3.2.2 Emotions	183
5.4 THE APPLICABILITY OF THE LTP MODEL	186
5.5 FINDINGS FROM THE QUANTITATIVE ANALYSIS OF THE THINK ALOUD DATA	188
5.5.1 Frequency of Test-taking Strategies as Reported by Test-takers	188
5.5.2 Frequency of Emotions, Emotional Regulation Processes and Behaviors as Reported by Test-takers	192
5.6 FINDINGS FROM THE TTSQ DATA	195
5.6.1 Test-taking Strategies	195
5.6.1.1 Test-taking Strategy Use	195
5.6.1.2 Qualitative Data on Test-taking Strategy Use	198
5.6.1.3 Results of the Factor Analysis of Test-taking Strategies	199

	<u>PAGE</u>
5.6.2 Emotional Regulation Processes, Emotions and Behaviors	207
5.6.2.1 Occurrence of Processes / Emotions / Behaviors	207
5.6.2.2 Qualitative Data on Emotions	210
5.6.2.3 Results of the Factor Analysis of Emotional Regulation Processes / Emotions / Behaviors	215
5.7 TEST-TAKING STRATEGIES AND EMOTIONAL REGULATION PROCESSES IDENTIFIED	220
5.8 SUMMARY OF FINDINGS AND OUTCOMES ON RESEARCH QUESTION 2	222
 CHAPTER 6: RESULTS & FINDINGS FOR RESEARCH QUESTION 3	 225
6.1 RESEARCH QUESTION 3	225
6.2 PROFICIENCY PROFILE OF RESPONDENTS	225
6.3 DO TEST-TAKING STRATEGIES VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?	227
6.4 IF SO, HOW DO TEST-TAKING STRATEGIES VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?	228
6.5 DOES EMOTIONAL REGULATION VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?	229
6.6 IF SO, HOW DOES EMOTIONAL REGULATION VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?	230
6.7 IS THERE A RELATIONSHIP BETWEEN TEST-TAKING STRATEGIES AND EMOTIONAL REGULATION?	231
6.7.1 Discriminant Analysis by Anxiety	235
6.7.2 Discriminant Analysis by Confidence	237

	<u>PAGE</u>
6.8 IF SO, DOES THE RELATIONSHIP BETWEEN TEST TAKING STRATEGIES SELECTED AND EMOTIONAL REGULATION DIFFER ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?	239
6.8.1 Discriminant Analysis by Proficiency Level	242
6.8.2 Discriminant Analysis by Two Proficiency Levels: Novice and Advanced	246
6.9 TEST-TAKING STRATEGIES, EMOTIONAL REGULATION AND TEST PERFORMANCE	248
6.10 SUMMARY OF FINDINGS AND OUTCOMES ON RESEARCH QUESTION 3	249
6.10.1 Research Question 3.1	249
6.10.2 Research Question 3.2	249
6.10.3 Research Question 3.3	250
6.10.4 Research Question 3.4	250
6.10.5 Research Question 3.5	250
6.10.6 Research Question 3.6	251
 CHAPTER 7: CONCLUSION	 252
7.1 CONCLUSIONS FROM FINDINGS	252
7.1.1 Defining a Context-Specific Construct of Language Proficiency	252
7.1.2 Test-taking Strategies and Emotional Regulation in a Test-taking Context	253
7.1.3 Test-taking Strategies, Emotional Regulation and Test Performance	254

	<u>PAGE</u>
7.2 IMPLICATIONS OF THE STUDY	255
7.2.1 Implications for Practice: Defining Context-Specific Constructs of Language Proficiency	255
7.2.2 The Relationship between Test-taking Strategies and Emotional Regulation	256
7.2.3 The Relationship between Test-taking Strategies, Emotional Regulation and Test Performance	257
7.2.4 The Language Testing Processing (LTP) Model	258
7.2.5 Learner Awareness of Test-taking Strategies and Emotional Regulation	262
7.2.6 Test-taking Strategy and Emotional Regulation Training	263
7.2.7 Implications for Test Developers	264
7.2.8 Methodological Implications	265
7.3 LIMITATIONS OF THE STUDY	265
7.4 RECOMMENDATIONS FOR FURTHER RESEARCH	266
7.5 SUMMARY OF MAJOR FINDINGS	268
 REFERENCES	 269
APPENDICES	282

LIST OF TABLES

	<u>PAGE</u>
CHAPTER 2: LITERATURE REVIEW	
Table (1): Matrix of models of language proficiency in language testing	27
Table (2): Summary of test-taking strategy research studies	43-45
Table (3): The social-cognitive model of self-regulation	58
CHAPTER 3: RESEARCH DESIGN	
Table (4): Comparison of positivism, pragmatism and interpretivism	76
Table (5): List of test-taking strategies obtained from the literature and included in the TTSQ	94
Table (6): Emotions, emotional regulation processes and behaviors obtained from the think aloud data and the literature	96
Table (7): Semi-structured interview respondents	100
Table (8): Respondents in the think aloud protocols	101
Table (9): Evaluation of the research design of the thesis	116-117
Table (10): Codes for different contexts of language use	122
Table (11): Codes for components of language proficiency	122
Table (12): Test-taking strategies obtained from the think aloud data	124
Table (13): List of codes to analyze the think aloud data: Test-taking strategies	125-126
Table (14): List of codes to analyze the think aloud data: Emotional regulation processes	127
CHAPTER 4: RESULTS & FINDINGS FOR RESEARCH QUESTION 1	
Table (15): Contexts of language use identified by students and teachers	132

	<u>PAGE</u>
Table (16): Components of language proficiency identified by students and teachers	136
 CHAPTER 5: RESULTS & FINDINGS FOR RESEARCH QUESTION 2	
Table (17): Affective responses of Test-taker A	146
Table (18): Affective responses of Test-taker B	150
Table (19): Affective responses of Test-taker C	154
Table (20): Affective responses of Test-taker D	157
Table (21): Affective responses of Test-taker E	160
Table (22): Affective responses of Test-taker F	163-164
Table (23): Affective responses of Test-taker G	167-168
Table (24): Affective responses of Test-taker H	170-171
Table (25): Affective responses of Test-taker I	172
Table (26): Affective responses of Test-taker J	175
Table (27): Affective responses of Test-taker K	177
Table (28): Affective responses of Test-taker L	178
Table (29): Test-taking strategies identified from the think aloud data	179-181
Table (30): Emotions identified from the think aloud data	184
Table (31): Frequency count of test-taking strategies reported	188-189
Table (32): Frequency count of emotional regulation processes / emotions / behaviors reported	192-193
Table (33): Means and SD of the test-taking strategies in the TTSQ	195-196
Table (34): Frequency distribution of TTS3 and TTS11 by percentage (N=497)	197
Table (35): Factor analysis for test-taking strategies	200
Table (36): Item loadings on factor 1 (TTS)	200
Table (37): Item loadings on factor 2 (TTS)	201
Table (38): Item loadings on factor 3 (TTS)	202

	<u>PAGE</u>
Table (39): Item loadings on factor 4 (TTS)	202
Table (40): Item loadings on factor 6 (TTS)	203
Table (41): Item loadings on factor 7 (TTS)	203
Table (42): Item loadings on factor 8 (TTS)	204
Table (43): Item loadings on factor 9 (TTS)	204
Table (44): Item loadings on factor 11 (TTS)	205
Table (45): Item loadings on factor 12 (TTS)	205
Table (46): Item loadings on factor 13 (TTS)	205
Table (47): Item loadings on factor 14 (TTS)	206
Table (48): Item loadings on factor 15 (TTS)	206
Table (49): Means & SD of the emotional regulation processes, emotions and behaviors in the TTSQ	207-208
Table (50): Frequency counts of coded qualitative data on respondents' feelings (N=351)	210
Table (51): Frequency counts of coded qualitative data on anxiety experienced by test section (N=314)	213
Table (52): Factor analysis for emotional regulation processes / emotions / behaviors	215
Table (53): Item loadings on factor 1	215
Table (54): Item loadings on factor 2	216
Table (55): Item loadings on factor 3	217
Table (56): Item loadings on factor 4	217
Table (57): Item loadings on factor 5	217
Table (58): Item loadings on factor 6	218
Table (59): Item loadings on factor 7	218
Table (60): Item loadings on factor 8	219
Table (61): Test-taking strategies and emotional regulation processes / emotions / behaviors experienced by Egyptian adult EFL learners	224

CHAPTER 6: RESULTS & FINDINGS FOR RESEARCH QUESTION 3

Table (62):	Distribution of respondents by proficiency level (N=497)	226
Table (63):	Means and SD of test-taking strategies by proficiency level	227
Table (64):	Means and SD of emotional regulation processes by proficiency level	229
Table (65):	Significant correlations between test-taking strategies and emotional regulation processes / emotions / behaviors	231
Table (66):	Distribution of high anxiety and low anxiety test-takers by proficiency level	234
Table (67):	Distribution of high confidence and low confidence test takers by proficiency level	234
Table (68):	Results of discriminant analysis by anxiety	235
Table (69):	Jackknifed classification matrix for high / low anxiety groups	235
Table (70):	Standardized canonical discriminant coefficients for anxiety	236
Table (71):	Results of discriminant analysis by confidence	237
Table (72):	Jackknifed classification matrix for high / low confidence groups	237
Table (73):	Standardized canonical discriminant coefficients for confidence	238
Table (74):	Significant correlations between test-taking strategies and emotional regulation processes / emotions / behaviors by proficiency level	239-240
Table (75):	Results of discriminant analysis by proficiency level	243
Table (76):	Jackknifed classification matrix for the four proficiency groups	243
Table (77):	Standardized canonical discriminant coefficients for proficiency level	244

	<u>PAGE</u>
Table (78): Results of discriminant analysis by two proficiency levels	246
Table (79): Jackknifed classification matrix for novice / advanced proficiency groups	246
Table (80): Standardized canonical discriminant coefficients for novice and advanced proficiency groups	247

LIST OF FIGURES

PAGE

CHAPTER 2: LITERATURE REVIEW

Figure (1):	Messick's view of validity	9
Figure (2):	Metacognitive strategies in language test performance	23
Figure (3):	The social-cognitive model of self-regulation	58
Figure (4):	The model of adaptable learning	62
Figure (5):	The proposed language testing processing (LTP) model	70

CHAPTER 3: RESEARCH DESIGN

Figure (6):	Research design	80
-------------	-----------------	----

CHAPTER 7: CONCLUSION

Figure (7):	The final language testing processing (LTP) model	261
-------------	---	-----

LIST OF APPENDICES

	<u>PAGE</u>
Appendix 1: Lists of Test-taking Strategies Identified from the Literature	282
Appendix 2: Semi-Structured Interview: Student Interview Protocol	288
Appendix 3: Open-Ended Questionnaire: Teachers' Questionnaire	289
Appendix 4: Likert-Type Questionnaire: Test-taking Strategies Questionnaire (TTSQ)	290
Appendix 5: Transcripts of the Pilot Student Interview Protocol	300
Appendix 6: Preliminary Version of the Test-taking Strategies Questionnaire	303
Appendix 7: Record of the 36 Student Interview Protocols	306
Appendix 8: Teachers' Questionnaire Qualitative Data	320
Appendix 9: Instructions and Example Items in the EPT (English language Placement Test)	326
Appendix 10: Four Sample Transcripts of the Think Aloud Protocols	333
Appendix 11: Four Sample Records of Think Aloud Data on Item Responses	348
Appendix 12: Miles & Huberman's (1994) Proposed Framework for Evaluation of the Research Design	352
Appendix 13: Four Samples of the Completed TTSQ	355
Appendix 14: BERA (British Educational Research Association) Guidelines on Ethics in Conducting Research	367
Appendix 15: Scree Plots of the Factor Analyses	371
Appendix 16: Descriptions of the Four Proficiency Levels at the Center for Adult & Continuing Education at The American University in Cairo	372

CHAPTER 1:

INTRODUCTION

This chapter describes the background of this study and the research questions to be addressed. An overview of the whole thesis is provided by briefly describing the contents of the following chapters.

1.1 BACKGROUND OF THE STUDY

English language is stressed in Egyptian education at all levels. It is taught as a foreign language in governmental schools starting at grade one and as a second language starting in kindergarten in private language schools. English is also the medium of instruction in many universities including the colleges of medicine, engineering, science and agriculture. Furthermore, English language proficiency is a requirement for students to register for graduate studies in many university programs. English ability is associated with educational achievement, which in turn determines social status. Prestigious professions require a certain level of proficiency in English and career advancement in Egypt in many fields is affected by the ability to communicate fluently in English. However, in spite of the fact that English is an integral component of the Egyptian school curriculum and that, across the board, access to public education in Egypt is theoretically free, thousands of adults enroll annually in EFL (English as a Foreign Language) evening classes. There are two possible explanations for this phenomenon. The first one is that these adults have not acquired the language skills they were supposed to learn through the formal education system. A second is that these adults recognize their own inadequacy in English language proficiency or perceive themselves as lacking in language skills, and thus, place themselves in a classroom learning situation with specific needs and expectations. Thus, the decision by an adult Egyptian to enroll in an EFL class may be seen as a form of self-assessment on the part of the learner.

Enrolment in these adult EFL programs is usually based on the results of an English language placement test. Placement tests used are either ready-made commercially available tests such as the institutional TOEFL (Test of English as a

Foreign Language) or tests designed in-house for the program. Placement tests are usually in multiple-choice format and may include a writing component. This test format is typical of many language proficiency tests in large-scale testing contexts whose sheer size imposes a number of constraints on the testing process including the limits on the time and money allocated for a language test which, in many cases, tends to pressurize testers to test what is easy or cheap to test. Such constraints increase the gap between the English used in real-world situations in a society and the English experienced in a testing situation (Hill & Parry, 1994). Accordingly, in many cases, scores on such tests neither reflect the learners' ability to use English in realistic workplace / academic contexts nor do they match the learners' own self perceived language ability.

The English Language Division of the Center for Adult & Continuing Education (CACE) at the American University in Cairo is the largest EFL program in Egypt. It has branches extending all over Egypt and approximately 4000 students enroll each term. The English language placement test of CACE consisted of a grammar multiple-choice test and a writing component.

This study was stimulated by both practical and theoretical considerations in the field of language testing in this particular context. While dealing with applicants enrolling in the EFL program, the researcher observed that many students after receiving the results of the test would ask to be replaced into a lower level without even attending the first class. They would indicate that the test results did not match their own assessment of language ability. There were fewer students requesting to be placed in a higher level. On analyzing the placement test it was found that it had been constructed on the basis of a traditional view of language: grammar and writing and there was no indication that students' needs had been taken into account when designing the test. The researcher's interest in investigating learners' perceptions of proficiency was prompted by the lack of a definition of English language proficiency which should form the basis of EFL curriculum and test design. Having observed the impact of self-

assessment on subsequent placement decisions, the researcher was also interested in exploring the cognitive and affective factors that affect language test performance.

1.2 PURPOSE OF THE STUDY

The aim of the present study is threefold. First, it attempts to establish a contextualized definition of language proficiency in an EFL context at CACE, AUC. Second, it seeks to identify learners' test-taking strategies and emotional regulation processes used during test-taking. Third, it seeks to investigate the relationship between test-taking strategies, affective factors and performance on a language test. Such insights on test-taking strategies and affective factors that influence test performance will also provide valuable information for curriculum designers and teachers to be incorporated into their classrooms, teaching activities and materials.

1.3 RESEARCH QUESTIONS

The specific research questions addressed in this study are:

1.3.1 Research Question 1

How is the construct of foreign language proficiency defined in a specific Egyptian EFL context?

- 1.1 What are the different contexts of language use of Egyptian adult EFL learners?
- 1.2 How do adult EFL learners in a specific context in Egypt define language proficiency?
- 1.3 To what extent does the construct of language proficiency vary according to the context of language use?

1.3.2 Research Question 2

What are the test-taking strategies and emotional regulation processes used and emotions experienced by adult learners in a specific Egyptian EFL context when taking an English language placement test?

1.3.3 Research Question 3

What is the relationship between test-taking strategies, emotional regulation and test performance for Egyptian adult EFL learners in a specific context when taking an English language placement test? Further questions that focus on the relationship between test-taking strategies, emotional regulation and test performance are:

- 3.1 Do test-taking strategies vary across different levels of test performance?
- 3.2 If so, how do test-taking strategies differ across different levels of test performance?
- 3.3 Does emotional regulation vary across different level of test performance?
- 3.4 If so, how does emotional regulation differ across different levels of test performance?
- 3.5 Is there a relationship between test-taking strategies selected and emotional regulation?
- 3.6 If so, does the relationship between test-taking strategies selected and emotional regulation differ across different levels of test performance?

1.4 OVERVIEW OF THE THESIS

Chapter 2 reviews the literature in several areas relevant to the research questions. The areas covered by the literature review are: language testing models, language test-taking strategies, language self-assessment, research in affect and test performance including language anxiety, self-regulation and emotional regulation during test-taking. The research questions are explicitly linked to the literature. The chapter also describes the Language Testing Processing (LTP) model derived from the literature.

Chapter 3 describes the research design of this study in detail and the research paradigm of where the study is located. The respondents, design and piloting of the data collections instruments, the data collection procedures and the analysis of qualitative and quantitative data are also described in detail. The research design is evaluated demonstrating the quality of the research and the validity and reliability of the findings.

Chapters 4, 5 and 6 present the findings on each of the research questions. In Chapter 4 the results related to defining the construct of language proficiency in a specific Egyptian EFL context are analyzed and discussed. In Chapter 5 the findings related to the test-taking strategies and emotional regulation of Egyptian adult learners in a specific context are discussed. In Chapter 6, the findings on the relationship between test-taking strategies, emotional regulation and test performance are analyzed and discussed.

Chapter 7 discusses the implications and conclusions of the study. Limitations of the study and recommendations for further research are also examined. A summary of the major findings of the study is provided.

CHAPTER 2:

LITERATURE REVIEW

The purpose of this chapter is to review the literature relevant to the research questions and to demonstrate how the three research questions emerged as gaps in the literature that needed to be addressed. In relation to Research Question 1, the area of language testing models was reviewed. The following areas were covered by the literature review in relation to Research Questions 2 & 3: language test-taking strategies, language self-assessment, research in affect and test performance including language anxiety, self-regulation and emotional regulation during test-taking. This chapter also describes a proposed Language Testing Processing (LTP) model which is based on the modification of the metacognitive strategies component of the Bachman & Palmer (1996) model, incorporating elements derived from the literature on language test-taking strategies, self-regulation and emotional regulation during test-taking.

2.1 VALIDITY & LANGUAGE TESTING MODELS

The two most important concerns in language testing are validity and the construct of language proficiency. Developments in the past twenty years in language testing have resulted in a profound change in the understanding of these two concepts. Theoretical advances in language testing emphasize the social dimension of validation research and challenge the traditional view of “language ability as consisting of skills (listening, speaking, reading and writing) and components (e.g. grammar, vocabulary, pronunciation)” (Bachman, 2000:2). Research into the construct of language proficiency now emphasizes the socially constructed nature of language use. Validity is the dominant theme of research in language testing and the overriding concern in test design is to demonstrate the validity of the interpretations and use of test scores. Thus, since the 1960s the focus of language testing research has been validation. Validation is a continuous process that ensures the appropriate use of tests. There is no one test that is valid for all time but a test has to be continually revisited and examined. Furthermore, there is no single trait underlying language proficiency that can be described in general but

language use must be determined for each specific context (McNamara, 2001; Bachman, 2000).

2.1.1 Importance of Validity

Messick (1995: 5) defines validity as an “overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores or other modes of assessment.” Validity is a property of the meaning of the test scores and applies to all types of assessment. Cumming (1996: 1) stresses the importance of validity and states:

“Validation in language assessment is ominously important, arbitrating educational and linguistic policies, institutional decisions, pedagogical practices, as well as tenets of language theory and research. But establishing validity in language assessment is by all accounts problematic, conceptually challenging, and difficult to achieve - probably more so than is recognized outside the specialized spheres of ...”

2.1.2 Changing Concepts of Validity

In the early beginnings of standardized testing, validity was determined by different procedures using diverse external criterion measures, item analysis and factor analysis. Different test authors used a confusing array of names for validity including face validity, intrinsic validity, logical validity, concurrent validity, predictive validity and empirical validity to name a few (Anastasi, 1986). In 1954, in an effort to standardize test construction procedures, the American Psychological Association, the American Educational Research Association and the National Council on Measurement in Education published standards for educational and psychological testing. This publication classified validity into four categories: content, predictive, concurrent and construct validity. In subsequent revisions between 1966 and 1985, predictive and concurrent validity were subsumed under criterion validity and this tripartite classification of validity has continued until the standards were revised in 1999 (Anastasi, 1988; Moss, 1992). This tripartite division, although initially useful in clarifying validity, did have an adverse effect on testing. These three labels were first “regarded as three distinct types of validity and later as three essential aspects or components of validity” (Anastasi, 1988:2).

Messick argued that these were not logically distinct categories of equal importance and that construct validity subsumed content and criterion validity. He considered content and criterion validity as relevant but insufficient evidence for establishing the validity of a test.

Although in the 1985 version of the Standards validity is viewed as a “unitary concept requiring multiple types of evidence to support specific inferences made from test scores” (Moss, 1992: 232), however, the traditional tripartite classification was retained. Finally, the 1999 revised Standards retreated from the three-part conception of validity and adopted a constructivist view as the fundamental unifying framework for all validity.

2.1.3 Unified Validity Concept

In 1989, Messick presented a unified framework for assessment validation (see Figure 1) which illustrated the centrality of construct validity and introduced constructs of social and cultural differences including considerations of social consequences of test use. In order to justify an interpretation of a test score for a particular purpose we must present evidence of construct validity and the value implications of interpreting this score. In order to justify the use of scores for a specific purpose we must collect evidence supporting construct validity and relevance of the particular use of this score and we must also consider the consequences of the decision made on the basis of this test score. Messick offered a new perspective to validity and extended it beyond the positivist paradigm. He considered validation as scientific inquiry however, he believed that scientific theories and methods are not value neutral. Attention to social consequences means that any assessment information should be interpreted in light of the context and the individuals involved. Anastasi (1990) and Cronbach (1989) have also discussed the importance of context in test validation and use of test scores. Anastasi (1990: 483-484) argues that it is important to “identify the locus and range of cultural (or other experiential) context for which any given test is appropriate, and then to keep both the use of the test and the interpretation of its scores within those contextual boundaries”.

Messick contributed a more detailed way of thinking about the data needed to support a validity conclusion.

"In Messick's view, the interpretation from our assessments are intended to be generalizable across individuals and contexts with validity defined as the extent to which those generalizations are true. That is, validity is the degree to which the "predetermined" meanings of the score points of our assessment system can be "matched" to particular performances.."

(Hamp-Lyons & Lynch, 1998:256)

From an interpretive perspective, generalization is possible but it is constructed differently. Meanings cannot be predetermined but must be interpreted within their particular assessment context. These interpretations lead to generalized understandings which in turn influence future interpretations. Therefore, construct validity of a language test depends on how the construct “language ability” is defined in a specific context.

Figure (1): Messick’s view of validity

	Test Interpretation	Test Use
Evidential basis	Construct validity	Construct validity + Relevance / utility
Consequential basis	Construct validity + Value implications	Construct validity + Relevance / utility + Value implications + Social consequences

After Messick (1989)

Messick's and Moss's (1996) perspectives on validity suggest a potential shift in language assessment research away from a field dominated by the positivist approach. Messick’s (1989) framework (see Figure 1) has been the basis for most validation research to date, however, not all aspects have been sufficiently researched and there are gaps that require more attention.

2.1.4 Validation Research in Language Testing

In 1998, Kunnan surveyed language testing validation research in the post-1980 period and identified key research themes within Messick’s framework as follows: proficiency components, test dimensionality, test validation process, test

development, value system differences, test-taking process, test-taking strategies, test-taker characteristics, social consequences and washback, ethics, standards and equity, and alternatives. Kunnan (1998) then organized the research studies under each research theme. He examined the research conducted in the area of test-taking strategies and found only five studies (Nevo, 1989; Anderson et al, 1991; Allan, 1995; Wijgh, 1996; Purpura, 1997) investigating test-taking strategies. These studies investigated test-taking strategies of Hebrew, Spanish, Cantonese, Dutch, Czech and Turkish speakers in reading comprehension and grammar and none in the Egyptian or Arab context. Therefore, more research needs to be done in order to better understand test-taking strategies of test-takers in different EFL contexts such as Egypt.

The two research themes identified by Kunnan (1998) that were found to include affective factors in their investigation of language test design and test performance were value system differences and alternatives. Five studies (Cohen, 1984; Zeidner & Benoussan, 1988; Bradshaw, 1990; Brown, 1993; & Peirce & Stein, 1995) were classified under value system differences and focused on students' attitudes and reactions towards specific tests and test formats and their use in test development. Other research that was related to affective factors were studies in the area of self-assessment (Oscarson, 1978,1989; Davidson & Henning, 1985; LeBlanc & Painchaud, 1985; Bachman & Palmer, 1989; Blanche & Merino, 1989; Janssen-van Dielen, 1989; Blanche, 1990; Heilenman, 1990; Peirce et al, 1993; Moritz, 1995). None of these studies has led to a conceptual framework or model that includes affective factors in test development. Most of the research has focused on self-assessment as an alternative to typical tests. Very few studies have been conducted on investigating the processes underlying self-assessment (Moritz, 1995; Peirce et al, 1993; Heilenman, 1991) and much of the research lacks well-defined conceptual frameworks.

In spite of the changes that have occurred in learning theories and new trends that have emerged in language learning and teaching, until recently, language testing researchers have resisted incorporating affective and volitional factors into the theory of language assessment. Boekaerts (1991) states that both objective and subjective feelings of competence affect academic performance and achievement. In any learning context "evaluation of competence is pervasive and continuous" and so "assessment and self-assessment are central, unavoidable and highly informative

components of daily functioning” (Boekaerts, 1991: 2). Therefore, learners’ feelings and perceptions of their own abilities and skills are important when evaluating performance. Of all the models explored (see matrix on page 25), only a few language testers (notably Bachman, 1990; Bachman & Palmer, 1996) have proposed models of language ability which include learners’ cognitive abilities and affective schemata. However, Bachman & Palmer’s model (see discussion of model on page 21) still does not clearly describe the relationship between affective factors and test performance. Research in emotional regulation during test-taking (Schutz & DeCuir, 2002; Schutz & Davis, 2000; Zeidner, 1995) has shown that test performance is affected by the emotions or feelings that occur and that emotional regulation can be controlled.

In language learning, recent studies have focused on the interaction of cognitive and affective factors that affect learning (Ehrman, 1998; Dickinson, 1995 and Aoki, 1999). According to Ehrman (1998) although little learning is likely to place without cognitive processing, teachers have seen how learning can be profoundly affected by feelings. This is because cognitive functions can be disrupted or enhanced by feelings. “These influential feelings include motivation, stress, anxiety, response to threat, self-efficacy, and well-being” (Ehrman, 1998: 2). For example, research has shown that motivated learners achieve better than unmotivated ones, when all else is equal, or learners who experience high anxiety are unable to participate in oral language learning activities.

Historically emotions or affect have been seen as being in conflict with reason. Early philosophers emphasized the pursuit of reason, however they could not negate or ignore emotions since human beings do not function solely on reason. This perceived conflict between reason and affect led to early philosophers coining the “metaphor of master and slave to describe the relation between reason and emotions (p.127).” Thus, emotions were regarded as inferior and more primitive than reason. This uneasy relationship between reason or cognition and affect provides the background history from which current views about affect emerge (Schutz & DeCuir, 2002).

The field of language testing is now being challenged by the changes that have occurred in language learning and teaching and to date only two models in language testing, namely Bachman’s 1990 model and Bachman and Palmer’s 1996 model, included learners’ cognitive abilities and affective schemata.

2.1.5 Models in Language Testing

Language testing research has tended to focus on investigating components of language proficiency and on theories of language test performance, the rationale being that a clear definition of the construct of language ability would enable test designers to design tests that match the definition. The construction and the design of any language test is based on a theory or beliefs of what language is in general, what language proficiency consists of and what is involved in learning a language. Each theory underlying test design defines constructs (psychological concepts) and describes the relationship between these constructs. In order to make inferences about learners' language ability and accordingly make decisions, test designers must demonstrate how performance on the test reflects actual use of the language in real life in a specific context. Therefore, a conceptual framework is needed to demonstrate this relationship between language test performance and actual language use.

Several models or theoretical frameworks have been used for developing language tests. Since the late 1970s, language testers have put forward several models of language proficiency which have revealed a lack of consensus in their representation of proficiency (Chalhoub-Deville, 1997). Various researchers such as McNamara (1990) and Kenyon & Stansfield (1992) have shown that when assessing proficiency a variety of terms and components are used leading to the existence of diverse models. This diversity has been discussed by Alderson (1991: 8) who argues that

“the profusion of competing and contradictory models, often with very slim empirical foundations, inhibits the language tester or applied linguist from selecting “the best model” on which to base his or her language test.”

The following is a review of influential models used in developing tests reflecting the changes that have occurred in the field. Furthermore, with the exception of Bachman and Palmer's 1996 model, none of the models discussed included an affective component.

Two major movements have marked language testing theory in the past two decades (the 80s and 90s). The first movement led to the definition of levels or scales of proficiency that focus on describing the language produced by learners at various

stages evaluated against an idealized criterion: the educated native speaker (Chalhoub-Deville, 1997; de Jong & Verhoeven, 1992). The second movement is characterized by models of language proficiency attempting to describe the components of proficiency.

2.1.6 Scales of Proficiency

In search of a common frame for describing and assessing language proficiency in different contexts, several scales of language proficiency have been constructed. In the 1980s among the most influential scales were the ACTFL (American Council on the Teaching of Foreign Languages) Guidelines, 1986; the Interagency Language Roundtable, 1985 and the Australian Second Language Proficiency Ratings scale, 1984. These scales were seriously criticized by several researchers (Savignon, 1985; Bachman & Savignon, 1986; de Jong, 1987) regarding their “semi-structuralist approach to language proficiency and the sparseness of validation studies” (de Jong & Verhoeven, 1992:5).

The ACTFL guidelines is a widely implemented rating scale in academic settings in the United States. It consists of a 9-point scale ranging from zero competence to the level of an ‘educated native speaker’ with skill-level descriptions of listening, reading, writing and speaking. The ACTFL guidelines in particular have been the most criticized. There have been inconsistent findings in the literature with some studies supporting the ACTFL hierarchy (Dandonoli & Henning, 1990; Kenyon & Stansfield, 1992 and Kenyon, 1995) and others questioning its validity. Savignon (1985) and Bachman (1988) argued that there was no empirical evidence that the scale descriptions were calibrated and Brindley (1991) states that the data collection methods used to produce the performance descriptors were not made public. On the other hand, Omaggio (1986:14) states that ACTFL guidelines “reflect data collected from oral interviews carried out over the past 40 years...and describe how language learners typically function along a whole range of possible levels of competence”. Using Spolsky’s (1990) argument that there are three approaches to describing and assessing a foreign language: structural, functional and general, de Jong criticizes the guidelines stating that learners’ proficiency varied across the range of levels and the descriptions of each level should reflect these differences. He stated

“that at lower levels of proficiency learners are more dependent on structural features, subsequently acquire (some) functional skills, and that finally, at an advanced stage, learners acquire sufficient strategic competence ...which eventually leads to a general proficiency, where constituent subskills become undistinguishable”.

(de Jong & Verhoeven, 1992: 9)

Another criticism was related to the use of the ‘educated native speaker’ as a reference to measure non-native speaker proficiency. Chalhoub-Deville (1997: 9) states that “empirical evidence shows how native speakers vary in their perceptions of learners’ proficiency depending on their professional training and experiences”.

These rating scales in developing assessment tools were subject to debate and required further empirical evidence. Chalhoub-Deville (1997) stated that empirically validated rating scales are necessary for valid language assessment and Bachman & Palmer (1996: 211-212) argued that rating scales must be based on “theoretical definitions of the construct” and that they are “defined operationally in terms of criterion levels of ability”, ranging from “no evidence of this ability” to “evidence of mastery of the ability”. The number of levels is determined by the number of decision levels required. Determining the construct validity of these rating scales was difficult and this approach to defining language proficiency was problematic and open to criticism.

In 2001, the **Common European Framework of Reference for Languages: Learning, teaching, assessment** was published. It is a common reference for language learning, teaching and assessment to be used across Europe. It describes the knowledge and skills a learner needs to order to use a language for communication and it also defines six levels of proficiency. It is primarily geared towards adult learners. The Framework is taxonomic in nature and breaks down language competence into separate components. These competences focus on an interaction between receptive and interactive modes particularly for speaking. There are ‘can do’ descriptors for reception, interaction and production. In designing the framework a rigorous validation methodology was followed in order to avoid the criticisms of the previous rating scales designed.

“A systematic combination of intuitive, qualitative and quantitative methods was employed. First, the content of existing scales was analyzed in relation to categories of description used in the Framework. Then, in an intuitive phase, this material was edited, new descriptors were formulated, and the set discussed by experts. Next a variety of qualitative methods were used to check that teachers could relate to the descriptive categories selected, and that descriptors actually described the categories they were intended to describe. Finally, the best descriptors in the set were scaled using quantitative methods. The accuracy of this scaling has since been checked in replication studies.”

(Council of Europe, 2001:22)

Although there has been rigorous validation to define the construct, there are problems with this Framework and the whole competency movement approach. This competency framework is based on what test-takers can do rather than on what they know. Thus, the test-takers’ internal cognitive processes and characteristics that underlie performance are ignored. Performance on assessment tasks is far more complex than the behaviors described and validity research has exposed the unreliability of performance assessments resulting from variability in factors related to tasks, raters and the interaction between test-takers’ traits and context. Thus, the usefulness of assessment information from these performance assessments is limited especially in high stakes decisions. As McNamara (2001:342) states “powerful validity arguments against the competency movement have fallen on deaf ears.” Since competency frameworks such as the Common European Framework are politically mandated and reflect the purpose and the conceptual view of those who develop them, thus empirical “validity research that expose the limitations or even contradictions in framework-based constructs” have very little impact (Douglas, 2001; McNamara, 2001).

The following section describes the second approach to defining proficiency which discusses key models of language proficiency. There is a critical discussion of each model and at the end of the section a matrix is presented summarizing the main features and weaknesses of each model.

2.1.7 Models of Language Proficiency

2.1.7.1 Unitary Competence Hypothesis

The notion of overall language proficiency was originally derived from Carroll (1961) who argued that

“the high correlation obtained between various sections of TOEFL [Test of English as a Foreign Language] and other general tests of English suggests that in fact we might be dealing with a single factor, English proficiency”

(Spolsky, 1967: 38)

The 1980s started with the Oller Unitary Competence Hypothesis exerting a strong influence on language testing. Oller (1979) conducted a series of studies using factor analysis to analyze correlations between various language proficiency measures. Results of principal components analysis led to a single, powerful common first factor which Oller labelled “unitary language competence”. There is now considerable doubt regarding this statistical argument because this method which is designed to simplify data is expected to produce or exaggerate the size of a single factor. More importantly, this general proficiency factor does not explain the total variance. By using different techniques, it is possible to explain the remaining variance by further independent factors. Subsequent studies (Bachman & Palmer, 1981, Carroll, 1983) showed that Oller’s hypothesis was not supported and that his methodology was flawed. Oller (1984) acknowledged the criticism regarding the validity of his statistical arguments and withdrew his claim regarding a unitary trait hypothesis. Davies (1981) also pointed out that Oller’s concept of an underlying unitary competence contradicts considerable evidence of the existence of at least two competences: productive and receptive processing of language in a social context. Moller (1982) pointed out that although Oller advocated global integrative tests such as cloze and dictation which integrate disparate language skills, these tests do not require test-takers to perform relevant, authentic communicative tasks. Furthermore, these global tests do not provide the test-takers the opportunity for spontaneous production either orally or in writing. Therefore, the construct validity of these indirect measures of language ability which have high standards of reliability and concurrent validity remains suspect.

2.1.7.2 Munby’s Model

Another model that was influential in language proficiency testing was that of Munby’s Communicative Needs Processor in which “the appropriate specification of communicative competence is processed from a profile of language communication needs” (Munby, 1978). Thus, test designers decide on the content of the test by analyzing students’ language needs for specific target situations using Munby’s

model and construct items accordingly. The design of the British Council English Language Testing Service (ELTS) test which was first introduced in 1980 was based on this model. It consisted of separate reading, listening, writing and speaking tests in different subject areas related to students' different fields of study. It also included a study skills component (Alderson & Clapham, 1992). Another test applying the Munby model was the Associated Examining Board's Test of English for Educational Purposes designed by Weir (1983). Munby's model has been severely criticized by many researchers. Davies (1981) and Skehan (1984) argued that it was impractical, not comprehensive and not theoretically plausible. Language testers such as Clapham (1981) criticized the validity of the specifications of the ELTS test. They "point out that no empirical data was ever gathered on the basis of which one might have justified the claims" (Alderson & Clapham, 1992: 152).

Bachman & Palmer (1996) criticize the approach that describes language ability as consisting of four skills (listening, speaking, reading and writing) and several components such as vocabulary and grammar. Characterizing the four skills in terms of channel (audio, visual) and mode (productive, receptive) is considered inadequate for two reasons. First, it does not distinguish between different language activities within the same skill such as taking lecture notes and writing an essay. Second, this approach does not take into account that the use of language occurs in "specific situated language use tasks" (p.75). They "argue that the concept that has been called 'skill' can be much more usefully seen as a specific combination of language ability and task characteristics.... (setting, input, expected response and relationship between input and expected response)" (p.76).

2.1.7.3 Canale and Swain Model

The more inclusive model of Canale and Swain (1980), further elaborated by Canale in 1983, was very influential in language testing research. This model proposed four components of communicative competence:

1. grammatical or linguistic competence – knowledge of the language code (eg. vocabulary, morphological rules, syntactic rules, phonological rules, sentence formation, spelling, ...etc.);
2. sociolinguistic competence – ability to understand and appropriately use language in different sociolinguistic contexts (eg. register and politeness);

3. discourse competence – ability to combine meanings and forms into different types of cohesive or unified texts (eg. narrative essay, poetry);
4. strategic competence – mastery of verbal and nonverbal communication strategies to communicate effectively and where necessary, overcome breakdowns in communication.

Schachter (1990) critically analyzed the model and questioned the validity of separating discourse and sociolinguistic competences. She argued that

“unity of text involves appropriateness and depends on contextual factors such as status of participants, purpose of the interaction, and norms or conventions of interaction”

Schachter (1990: 43)

Another criticism was the unclear relationship between the four competences which led to ambiguous definition of theoretical constructs for test design. The model does not explain “the nature of the transition to communicative performance” (Skehan, 1991: 8). Empirical studies to validate the communicative competence model were conducted by Allen et al (1983) and Harley et al (1990) using factor analysis. They were not able to differentiate between grammatical, discourse and sociolinguistic competences. Bachman (1990) analyzed the methodology used and concluded that the model was not supported “owing to either a large method effect or problematic operationalization of the traits, ...” (Chalhoub-Deville, 1997: 6). In spite of the criticisms, this model was extremely influential in language testing during the 1980s.

2.1.7.4 Bachman Model

In 1990, a more comprehensive model of language test performance was proposed by Bachman. The model was developed based on both theory and state-of-the-art empirical and measurement techniques (Kunnan, 1995). It goes beyond the Canale and Swain framework and addresses the competence-performance relationship. The Canale and Swain linguistic and discourse competences are drawn together and included under organizational competence.

The Bachman (1990) model represents “a milestone in language testing” (Skehan, 1991: 15). None of the previous models explained performance on language tests and none included the various factors that influenced test

performance. The Bachman Model (1990) of communicative competence and performance does just that. Bachman states

“The four categories of influence on test scores included in this model are communicative language ability, test method facets, personal characteristics and random measurement error.”

Bachman (1990: 348)

In the model there are four factors that are seen to influence performance on a test: language ability, test methods, test-taker characteristics and random factors. The following is a brief description of each factor.

Bachman divided language ability into three components: language competence, strategic competence and psychophysiological mechanisms.

1. **Language competence** is defined as “a set of specific knowledge components that are utilized in communication via language” (Bachman, 1990: 84). It is divided as follows :
 - **Organizational Competence**
 - Grammatical competence (morphology, syntax, vocabulary)
 - Textual competence (cohesion and rhetorical organization)
 - **Pragmatic Competence**
 - Illocutionary competence (language functions)
 - Sociolinguistic competence (register, cultural allusions)
2. **Strategic competence** is defined as “the mental capacity for implementing the components of language competence in contextualized communicative language use” (Bachman, 1990: 84). It includes assessment, planning and execution.
3. **Psychophysiological mechanisms** are defined as “the neurological and psychological processes involved in the actual execution of language as a physical phenomenon” (Bachman, 1990: 84).

The second component of Bachman’s model is test method facets which include :

- characteristics of the test environment (place, equipment, time and physical conditions);
- characteristics of the test rubric (test organization in terms of sequence, time allocated, specification of procedures and tasks, scoring procedures;
- characteristics of the test input (format of test in terms of channel, mode, form and language of input);

- characteristics of expected response;
- restriction on response;
- relationship between input and response

It is interesting to note the inclusion of test methods, that is, ways in which the format of the test itself may intrude on or affect the measurement. This is an important aspect of the model because it implies that testers must examine the systematic effects of these factors since part of a test result may be due, in part, to test format effects rather than underlying language ability.

The third component in Bachman's model is test-taker characteristics which include cultural background, background knowledge, cognitive abilities, sex and age. Recent studies investigating these factors (Kunnan, 1995; Purpura, 1996) have shown that these factors can have a significant impact on language test performance.

The fourth category is random measurement error which consists of interactions among the three components of the model and measurement error. In order to ensure that language test scores reflect language ability and not other factors, test methods, test-taker characteristics and random factors need to be examined so as to minimize their effects on test performance (Kunnan, 1995).

McNamara (1996) notes that Bachman's model is superior to that of Canale & Swain for testing purposes because it attempts to separate knowledge of the language from ability to use the language which Bachman referred to as strategic competence. However, McNamara notes that there is still some overlap between illocutionary and strategic competences but he does not offer an alternative model of language ability. He only describes the problem but does not provide any possible solutions. He states that it is difficult to apply frameworks of language ability which include separate features of language use to practical test design because performance is not a cluster of these features but an interaction among them.

The Bachman model has been criticized as being difficult to apply in the actual implementation of tests and presents a problem for test designers (McNamara, 1990). Bachman (1990) himself viewed the model as a guide for researchers in language testing. Although Bachman's model was a useful tool, it still lacked a clear relationship between communicative language ability and test construction. Chalhoub-Deville (1997) also criticized the model for not clearly explaining the relationship between language proficiency as defined in the model and context-specific language constructs. Furthermore, the strategic competence component was

essentially a general cognitive capacity and did not include affective or volitional factors. In short, Bachman's model was still lacking in two key aspects: context-specific constructs of language ability and affective factors. These two factors were included in Bachman & Palmer's model described in the following section.

2.1.7.5 Bachman & Palmer's Model

Another model proposed by Bachman & Palmer (1996: 17) which is the latest version of the Bachman (1990) model is based on the premise that usefulness is the most important quality of a test. They "believe that test usefulness provides a kind of metric by which we can evaluate not only the tests that we develop and use, but also all aspects of test development and use." Test usefulness consists of six test qualities: reliability, construct validity, authenticity, interactiveness, impact and practicality. Usefulness is operationalized on the basis of three principles:

- Principle 1: The focus is on maximizing overall test usefulness and not the test qualities.
- Principle 2: The test qualities must be evaluated in terms of their collective effect on usefulness rather than separately.
- Principle 3: There are no general prescriptions for the appropriate balance of test qualities or minimum level. This must be determined for each testing situation. Usefulness is evaluated in specific testing situations.

Although the research questions of this study focus on two components of the model, namely, construct validity and interactiveness, the six key test qualities of the model are briefly described as follows:

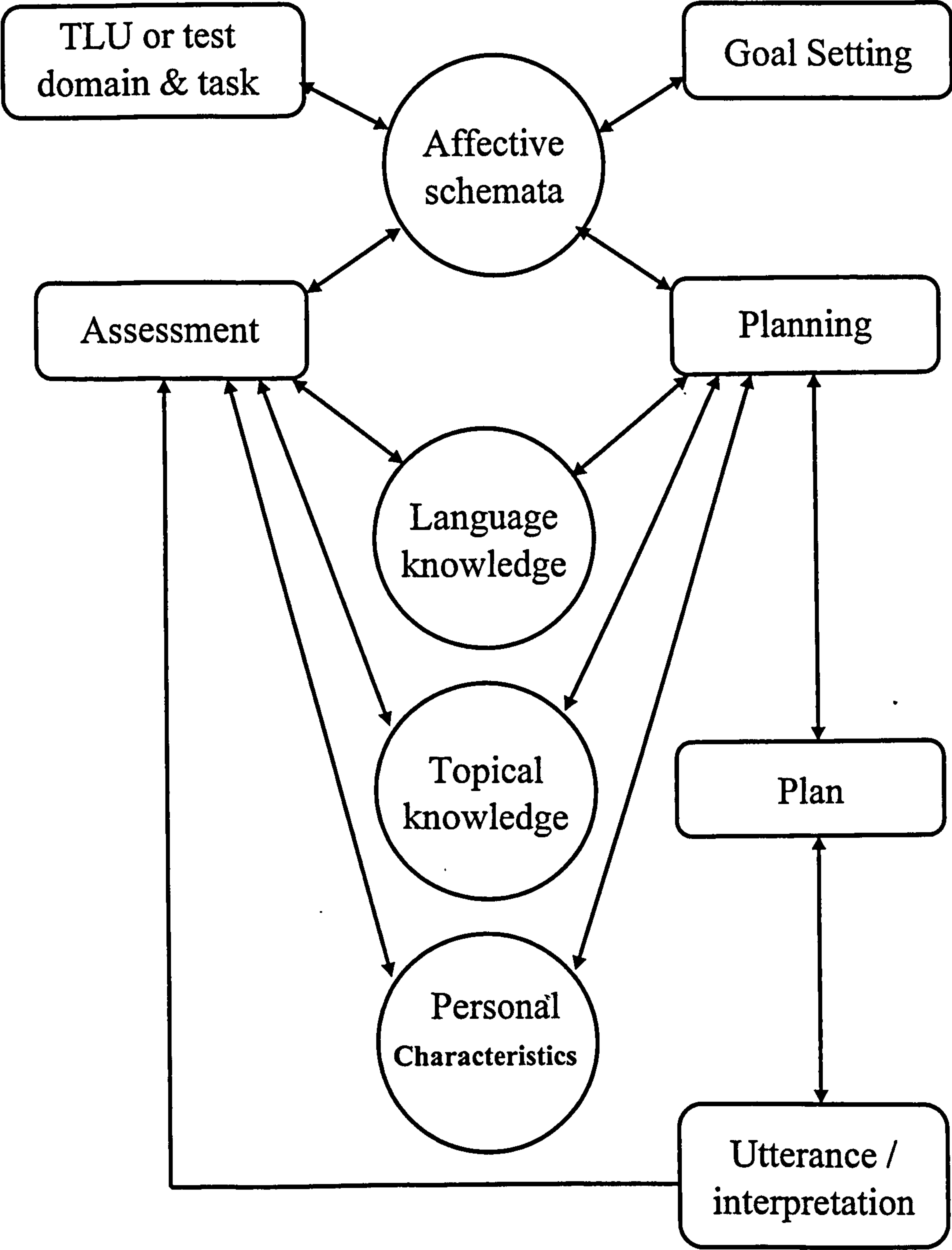
1. **Reliability** is "considered as a function of consistencies across different sets of test task characteristics" (Bachman & Palmer, 1996: 20).
2. **Construct validity** "refers to the extent to which we can interpret a given test score as an indicator of the ability(ies), or construct(s), we want to measure...with respect to a specific domain of generalization" (Bachman & Palmer, 1996: 21).
3. **Authenticity** is "defined as the degree of correspondence of the characteristics of a given language test task to the features of a target language

use (TLU) task” (Bachman & Palmer, 1996:23). They propose a framework of language task characteristics based on Bachman’s (1990) model which consists of five components : “setting, test rubric, input, expected response and relationship between input and response” (p.47).

4. **Interactiveness** is “defined as the extent and type of involvement of the test-taker’s individual characteristics in accomplishing a test task” (Bachman & Palmer, 1996: 25). They identify the following individual characteristics: language ability (language knowledge and strategic competence or metacognitive strategies), topical knowledge, and affective schemata. Their model of language ability is based on the Bachman (1990) model except for relabelling strategic competence as metacognitive strategies. Thus, language ability consists of language knowledge and metacognitive strategies which are goal setting, assessment and planning.
5. **Impact** or consequences of using tests occurs at two levels: the society and educational systems and individuals. Test developers and test users need to carefully analyze the results of using a test for a particular purpose. Washback has been an important issue on language testing and is considered an aspect of impact. Hughes (1989: 1) defines backwash (or washback) as “the effect of testing on teaching and learning”. Wall & Anderson (1993) note that washback affects not only individuals but the educational system as a whole. The impact of tests on test-takers, teachers, the society and educational system must be considered when designing tests.
6. **Practicality** is “defined as the relationship between the resources that will be required in the design, development and use of the test and the resources that will be available for these activities” (Bachman & Palmer, 1996: 39). It is related to whether or not the test will be designed or used because if the resources required exceed the resources available, then the test will be impractical. Resources are defined as human resources, material resources and time.

Figure (2) graphically illustrates Bachman & Palmer’s framework of the interaction between the characteristics of individuals which include language knowledge, metacognitive strategies, personal characteristics, topical knowledge, affective schemata; and characteristics of the test task or language use situation. These characteristics are viewed as possible influences on a test-taker’s performance on a language test. Personal characteristics include age, gender, nationality, resident status, native language, educational background and previous experience with tests. Topical knowledge is an individual’s knowledge of the world. Affective schemata consist of an individual’s emotions. Affective schemata combined with the

Figure (2): Metacognitive strategies in language test performance
(Bachman & Palmer, 1996: 72)



characteristics of a particular task determine an individual's affective or emotional response to the task. "In a language test, test-takers' affective schemata may influence the ways in which they process and attempt to complete the test tasks" (p.66) and thus test performance may be enhanced or inhibited. Bachman & Palmer state that this framework is not "a working model of language processing, but ... a conceptual basis for organizing our thinking about the test development process" (p.62).

They divide metacognitive strategies into three areas: goal setting, assessment and planning. Goal setting involves identifying the test tasks, choosing one or more tasks (where there is a choice) and deciding whether one can complete the task(s) or not. Assessment consists of assessing one's language and topical knowledge and relating it to the test task(s). This also involves activating an individual's affective schemata for handling the task. Assessment includes evaluating the correctness of one's response in relation to perceived criteria for completion of the task. Affective schemata are also involved in determining "the extent to which failure was due to inadequate effort, to the difficulty of the task, or to random sources of interference" (p.73). Thus, self-assessment is included as a component of the model and involves assessing the requirements for completing a test task, assessing one's abilities and assessing the correctness of the response to the test task. Planning involves deciding how to use elements from one's language knowledge, topical knowledge and affective schemata to complete a test task. The outcome of planning is a plan which when implemented is the actual response to the test task. It can be seen that an individual's affective schemata is involved in all areas of metacognitive strategy use. These goal-setting, assessment and planning strategies can also be viewed as test-taking strategies.

In addition to the above framework, Bachman & Palmer (1996) also included a set of procedures to be followed for designing 'useful' language tests. To ensure that usefulness is taken into account in all stages of test design, the test development procedures include a formal plan for evaluating usefulness. They propose a checklist of questions for evaluating usefulness and have specified "procedures for collecting qualitative and quantitative evidence (p.149). They present ten projects illustrating the actual application of their model to test development, changing test designs reflecting changes in testing situations, and developing classroom achievement tests.

This model is superior to the Bachman (1990) model since it has included practical procedures for test construction. Another strength is that the model does take into account the language construct of a given context. Bachman & Palmer differentiate between syllabus-based construct definitions and theory-based construct definitions. Test developers define constructs based on specific components of language ability in a course syllabus or on components described in a theory of language ability. They propose a language ability checklist to help test developers in defining the construct to be measured. This model reflects the field's current knowledge and clearly takes into account the paradigm shift that has occurred in language learning. This model provides a conceptual framework for language testing which takes into account new variables such as affective factors, metacognitive strategies, socially defined and context-embedded uses of language. They have also included self-assessment in their framework.

However, there are several criticisms of the model. McNamara (1996) notes that there is still some overlap between the language knowledge and metacognitive strategies components. The debate in language testing regarding the knowledge / skill balance or distinction is not new (Davies, 1996) and in the Bachman & Palmer (1996) model it is clearly indicated that metacognitive strategies are influenced by the testing event. However, the relationship between metacognitive strategies which include the affective schemata and test performance requires further clarification. For example, the model does not show how specific test-taking strategies such as guessing, skimming, ...etc would be included and it does not clearly explain the interaction between assessment and test performance. The model does not show the results of performance on a task on an individual's language and topical knowledge and assessment strategies.

Davies (1996) stated that on examining the actual use of the Bachman & Palmer model in test development / use projects, there seems "to be a suspicion of lip-service" (p.61). He went on to say that in a case study presented in Brindley (1995) the researcher had used the Bachman & Palmer framework as a general guide applying a 'weak' interpretation of the model. Thus, Davies concluded that it is "more like claiming to draw explicitly on current theoretical frameworks than actually doing so" (p.62). Another example is that although the model states that affective factors are to be taken into account during test design, however, this does not seem to be applied in current practice. In examples of test development projects,

affective factors do not appear in the actual test itself and are only mentioned in the specifications or design document in very general terms. An excerpt taken from one of the case studies regarding “possible affective responses to taking the test” reads as follows:

- “1 Highly proficient test-takers: likely to feel positive about taking the test, since it provides them with an opportunity to be exempted from the sheltered writing program, ...
- 2 Less proficient test-takers: may feel threatened by the test, since they may not meet the minimum standards for admission into the sheltered writing program.”

(Bachman & Palmer, 1996: 259)

It is not clear how these affective factors are reflected in the design of the test, the scoring procedures or decisions taken based on test results.

Another criticism is related to contextualized definitions of proficiency. North (1993) distinguished between theoretical models which describe proficiency at a general level and operational models which represent proficiency at the contextual level. Chalhoub-Deville (1997) asserts that test designers should construct assessment frameworks or operational models that are empirically developed within their contexts. She compared two assessment frameworks developed by Hinofotis, Bailey and Stern (1981) and Chalhoub-Deville (1995) to illustrate the differences reflecting their respective contexts. She added that “it is expected that the components that emerge using an empirical, contextualized approach should concur with the theoretical model that best represents the field’s current state of knowledge” (p.13).

Bachman & Palmer (1996) do stress the need for providing empirical evidence of the validity of the components of the construct being defined, however, they do not provide a framework for doing so. They also discuss the role of topical knowledge in defining the construct. Traditionally, language testers have considered topical knowledge to be a potential source of test bias or invalidity. Bachman & Palmer (1996) argue that this is not necessarily true since there are situations where it is part of the construct that a test developer may want to measure.

2.1.7.6 Summary of Models of Language Proficiency

The following matrix (on page 27) summarizes all the models previously discussed. The matrix illustrates how each of the five models defines language ability.

Table 1: Matrix of models of language proficiency in language testing

Model	Language Ability	Test Performance	Criticism(s) of Model
Oller's Unitary Competence Hypothesis (1979)	An underlying unitary competence		<ul style="list-style-type: none"> ◆ There was considerable evidence to indicate the existence of several competences underlying language ability.
Munby (1978)	Four skills : listening, speaking, reading, writing and vocabulary, grammar		<ul style="list-style-type: none"> ◆ It did not account for different abilities within same skill eg. taking notes & writing an essay. ◆ It did not account for language use in different situations.
Canale & Swain (1980)	4 components of communicative competence: <ul style="list-style-type: none"> ◆ grammatical / linguistic ◆ sociolinguistic ◆ discourse ◆ strategic 		<ul style="list-style-type: none"> ◆ It did not clearly describe the relationship between the 4 competences. ◆ The difference between discourse and sociolinguistic competence was not clear. ◆ Empirical studies to validate the model were not able to differentiate between grammatical, discourse & sociolinguistic competences.
Bachman (1990)	3 components: <ul style="list-style-type: none"> ◆ Language competence (organizational & pragmatic) ◆ Strategic competence ◆ Psychological mechanisms 	<ul style="list-style-type: none"> ◆ Test methods facets ◆ Test-taker characteristics ◆ Random measurement error 	<ul style="list-style-type: none"> ◆ It lacked a clear relationship between communicative language ability & test design. ◆ It did not explain the relationship between language proficiency & context-specific language constructs.
Bachman & Palmer (1996)	2 components: <ul style="list-style-type: none"> ◆ Language knowledge (organizational & pragmatic) ◆ Metacognitive strategies 	Test usefulness in specific testing situations consists of: reliability, construct validity, authenticity, interactiveness, impact & practicality	<ul style="list-style-type: none"> ◆ It lacks a framework for validating the language constructs being defined. ◆ It lacks a clear relationship between affective factors & test design ◆ It requires further development of the relationship between metacognitive strategies / affective factors and test performance

The matrix also shows that only the two of the most recent models that emerged in the 1990s included performance on a test. The matrix highlights the weaknesses in each model and shows the progressive development of the models with each model overcoming the weaknesses of the previous one. In conclusion, there is a variety of interpretations of the construct of English language proficiency in the literature depending on the model adopted and the context involved. There is no generic definition for English language proficiency which can be extended to this particular context in Egypt.

2.1.8 Research Question 1

Therefore, in order to apply the Bachman & Palmer model (1996) in a specific context, the construct of language proficiency needs to be empirically defined for that particular context. This leads the researcher to posit Research Question 1: How is the construct of foreign language proficiency defined in a specific Egyptian EFL context? (Research Question 1) This general question can be broken into three sub-questions focusing on specific contexts of language use, and on how adult EFL learners define language proficiency. Adults need English for a variety of purposes and uses and in order to define the construct as accurately as possible, these contexts must be identified. Furthermore, all adults have a view of what it means to be proficient in English. These views may be some theoretical concept of an idealized native speaker which may vary depending on their experiences and world knowledge. After all, native speakers do differ in their listening and speaking skills and there are individual differences in oral and written performance in terms of style and fluency. Thus, Egyptian adults may have different perceptions of English language proficiency. The three sub-questions are:

- What are the different contexts of language use of Egyptian adult EFL learners?
- How do adult EFL learners in a specific context in Egypt define language proficiency?
- To what extent does the construct of language proficiency vary according to the context of language use?

The literature review has also revealed that the relationship between affective factors, the goal-setting, assessment & planning strategies or test-taking strategies and performance on a test must be further elaborated or developed in the Bachman &

Palmer (1996) model. In order to explore affect and test-taking strategies in relation to test performance, three separate but related areas of research will be reviewed in the following sections: research in test-taking strategies, research in self-assessment and research in self-regulation. The aim of this review is to summarize the findings to be used as a basis to propose a language testing model based on the Bachman & Palmer model.

2.2 RESEARCH IN TEST-TAKING STRATEGIES

Since the late 1970s, interest in a process approach to test-taking has slowly begun to grow and researchers (Cohen & Aphek, 1979; Cohen, 1984; Anderson et al, 1991; Nevo, 1989) are investigating the strategies used by language test-takers while taking a test. The data obtained on test-taking processes are being used to validate language tests, however more research is required in this area. As Cohen (1998:108) states:

"Even though the field of test-taking strategy research is a fledgling one ... Consideration of the findings from this growing research area will undoubtedly prove beneficial at all points in constructing, administering, and interpreting language tests."

2.2.1 Defining Test-taking Strategies

Test-taking strategies are defined as specific actions, behaviors or techniques that test-takers apply consciously or unconsciously while taking a test (Cohen, 1998; Purpura, 1999). Cohen (2000:129) notes that strategy use "implies an element of selection." The test-taking process is viewed as the cumulative effect of using test-taking strategies. "Strategies vary according to context" (Cohen, 1998:92) and different respondents use different strategies for different types of tests taken in a variety of situations.

Some respondents may get by doing a surface matching between the information in the passage, in the multiple-choice item stems and in the item alternatives without processing for meaning. Other respondents use shortcuts to answer questions. For example the respondent does not read the passage as required but immediately starts looking for the answers to the given reading comprehension questions. Response to a test item may require the use of a limited number of strategies and some items may require the use of several strategies. "It is best not to assume that any test-taking strategy is a good or a poor choice for a given task"

(p.93). This depends on the test-taker's use of the strategy for a specific task. Furthermore, a strategy may be successful for one test-taker but may not be as successful for another test-taker.

The result of using these various strategies is that the test-takers may "get items wrong for the right reasons or right for the wrong reasons" (Cohen, 1984:71). A good student may get the answer wrong based on "an adventurous inference" while a student of lower ability may get the answer right based on incorrect reasoning. Therefore, due to using specific test-taking strategies or due to poorly designed test items, "students may not be displaying a representative performance of their language" (p.71) ability.

The ability of learners or test-takers to use strategies has been referred to as strategic competence (Cohen, 1998). Bachman & Palmer's (1996) model includes strategic competence as a set of metacognitive strategies: goal-setting, assessment and planning. Purpura (1999:6) defines metacognitive strategy use "as a set of conscious or unconscious mental or behavioral activities which are directly or indirectly related to some specific stage of the overall process of language ... testing." Purpura views metacognitive strategies as self-management strategies that are used to plan, monitor and evaluate a testing event. Cognitive strategy use "is defined as a set of conscious or unconscious mental or behavioral activities or operations which are directly or indirectly related to the comprehending, storing or retrieval of information during language .. testing" (p.7). Purpura (1999) states that most strategy researchers (O'Malley & Chamot, 1990; Oxford, 1990; Wenden, 1991) agree that in second language use situations learners use both cognitive and affective strategies. Thus, in a test-taking situation metacognitive, cognitive and affective strategies are all applied. Bachman & Palmer's (1996) model shows that metacognitive strategies do interact with affect and language knowledge, however the model does not adequately operationalize affective schemata and does not clearly describe the interaction.

2.2.2 Identifying Test-taking Strategies

Self-report methods through a questionnaire or checklist or using introspective or retrospective verbal report techniques have been the main tools used in collecting data on test-taking strategies. These methods are based on the assumption that test-takers would reveal at least part of their strategies when they

think aloud or use a self-report questionnaire or checklist to report the strategies they believed they used when taking a test. Verbal reports have provided insight into how test-takers take tests and have helped determine the strategies used by test-takers while performing specific testtasks. For example, studies were conducted collecting verbal reports on students taking multiple-choice reading tests (Cohen, 1998). However as Cohen (2000:131) notes:

“As researchers are only beginning to develop means for collecting more than just anecdotal data on test-taking strategies, we cannot say that verbal report methods have already yielded numerous valuable insights. We can, however, assert that these insights are beginning to appear in ever more systematic patterns...”

The validity of self-report data has been questioned and the reliability and validity of verbal reports is further discussed in this study in the chapter on Research Design. An important point that needs to be mentioned is that the current methodology used in collecting verbal reports or verbal protocol analysis is different from the approach used in earlier research (Green, 1998). As Cohen (1998:95) states "Earlier work reported on approaches that involved at most a request of respondents to reflect back on the strategies they used in arriving at answers to a subtest or group of items, producing data of more questionable reliability and validity."

2.2.3 Studies on Test-taking strategies

In this section, different studies on test-taking strategies are reviewed and the different test-taking strategies used by test-takers in different language test contexts and in completing different language tasks are reported. Most of the test-taking strategy research has been conducted on reading comprehension tests. The researcher has only been able to find one study investigating test-taking strategies for a listening test and one study for taking a proficiency test. The aim of this section is to summarize the key findings to date on test-taking strategies in the literature in order to identify the gaps, if any. At the end of the section, a matrix summarizing the findings is provided.

2.2.3.1 Reading Comprehension Tests

The research conducted on strategies used while taking reading comprehension tests is discussed in terms of different types of tests: multiple-choice, cloze, summarization, and short answer. Detailed descriptions of the methods used in each study are provided in order to find out the approaches used.

Multiple-choice reading tests

In 1984, Cohen reported on several student course projects investigating test-taking strategies of students taking reading tests with multiple-choice question or cloze tests. Because these studies were based on student course work they were limited in scope and as Cohen (1998) stresses, more research needs to be conducted in order to determine the generalizability of the findings. Nevertheless, these studies did yield interesting results and are described in this section and in the following section on cloze reading tests.

Larson (1981 in Cohen, 1984) asked 40 ESL university students to describe how they responded to a 10-item multiple-choice test which consisted of a 400-word reading passage. Retrospective verbal reports were obtained. The researcher interviewed 17 students in groups of two or three within 24 hours after taking the test and met with the rest of the 23 students in groups of five or six, four days after the test. The strategies that the respondents reported they used were:

"(1) they stopped reading alternatives when they got to the one that seemed correct to them, (2) they matched material from the passage with material in the item stem and in the alternatives ... (3) they preferred a surface-structure reading of the test items to one that called for more in -depth reading and inferencing" (Cohen, 1998:98)

The students' overall approach was to either read the questions first or just read parts of the passage corresponding to the questions.

Israel (1982 in Cohen, 1984) investigated the role of topical background knowledge and vocabulary knowledge in answering multiple-choice questions on a reading passage. The key finding was that students seemed to be relying on their background knowledge rather than reading the passage carefully.

Cohen (1998) describes a study conducted by Gordon (1987) involving 30 tenth-grade EFL students at two reading ability levels (high and low). Students were asked to think aloud while answering open-ended and multiple-choice questions. The

results showed that the "answers to test questions did not necessarily reflect comprehension of the text" (p.100) and that low ability and high ability students used different test-taking strategies. High ability students used more global processing strategies such as predicting from context and using lexical and grammatical knowledge while low ability students tended to use more word-centered strategies processing information at the sentence or word level. Low ability learners were likely to translate word for word or focus on key words in the text and match them with words in the questions or copy words from the text.

Waxman & Padron (1987) investigated the cognitive reading strategies used by 82 third, fourth and fifth grade ESL Hispanic students and the effect of these strategies on performance on a reading test. The test was the reading comprehension section of the Stanford Diagnostic Reading Test administered twice with an interval of 4 months between the two test administrations. Immediately after the post test, students were asked to complete a 14-item, Likert-type reading strategies questionnaire. The questionnaire consisted of 7 strategies that were found to be negatively related to students' reading achievement and 7 strategies that were found to be positively related to students' reading achievement. The results indicated that two particular strategies: saying the main idea over and over and thinking about something else while reading were negatively related to achievement. The results also showed that lower achieving students used inappropriate cognitive strategies.

Thus, the studies reviewed so far indicated that high ability and low ability students use different strategies and that low ability students use inappropriate strategies. The implications are that with strategy training, low ability students may perform better on tests and therefore, more research is needed to identify these appropriate and inappropriate strategies. The studies used either think alouds, retrospective verbal reports or questionnaires to collect data. None of them combined these methods.

Nevo (1989) conducted an interesting study on 42 tenth-grade students studying French whose L1 was Hebrew investigating L1 and L2 reading test-taking strategies using a checklist. Students were given a test consisting of two reading passages in Hebrew and two passages in French with 5 multiple-choice questions on each passage in the same language. In both languages, there was an easy passage and a more difficult one. The test also included a strategy checklist which consisted of 15 strategies identified from the literature and based on the researcher's intuitions about

strategies test-takers were likely to use. The checklist described each strategy briefly and provided a code name for each. It also included an option for test-takers to include strategies they use that were not on the checklist. Students were required to answer each question and then to immediately record which strategy was most instrumental in answering the question and the second most instrumental strategy. Following each passage and questions, test-takers were asked to indicate the numbers of the two easiest and the two most difficult questions and to describe in writing the nature of the difficulty. After completing the test, test-takers were requested to complete a questionnaire covering general strategies used in the test as a whole. Students' responses were kept anonymous in order to ensure that students would report what they actually did rather than what they believed was expected of them. The findings indicated that test-takers were able to provide feedback on their strategy use after each item using the checklist included with the test. The results also showed that there was transfer of strategies from L1 to L2. The researcher identified strategies that contributed to the choice of the correct answer (contributory strategies): using background knowledge, returning to the passage, looking for the answer in chronological order, looking for clues in the text, ceasing search for correct alternative once reaching plausible choice, selecting an alternative by eliminating others, selecting an alternative because of a key word that appeared in the text and selecting an alternative because of a word that the test-taker associated with another word in L1. Non-contributory strategies were also identified as those which did not contribute to the choice of the correct answer. It was found that test-takers reported the use of more non-contributory strategies in L2 than in L1. The researcher found a significant relationship between difficulty of text and strategies in L2. There was more use of contributory strategies in the easier text than in the difficult one and the two types of questions that led to difficulty were items which were grammatically or semantically complex or items which required inferencing or prediction. Test-takers reported that they benefited from taking the test by becoming aware of what they did when taking a reading test. Once again contributory and non-contributory strategies were identified and these were related to the difficulty of the reading text.

In research based on self-report data, it is important to ensure that the data obtained is valid and not dependent on the particular instrument being used. Allan (1995) criticized Nevo's study stating that instrument effect was not mentioned as a threat to validity and conducted a study investigating the effect of an instrument on

students' self-reports. The subjects were 5 intact groups of first-year university ESL students (109 students) whose native language is Cantonese. Each group took a test that consisted of a reading passage with 11 multiple-choice questions and a checklist. The study was conducted over five consecutive weeks with each group being tested once. In the first four weeks the checklist was modified from week to week based on the findings of the previous administration while in week 5 the checklist was not given. In week 1, the complete checklist was given and strategy number 3 that was most frequently reported was 'returning to the passage'. In week 2, the sequence of strategies listed was reordered with the strategy of 'returning to the passage' being listed towards the end of the checklist. The week 2 responses were analyzed and the most frequently reported strategy was again 'returning to the passage', indicating that position did not have an effect. In week 3, this strategy 'returning to the passage' was deleted and the list was renumbered. Week 3 results showed that instead of selecting the open-ended 'other strategy', most test-takers selected another strategy from the modified checklist: 'clues in the text' indicating an instrument effect. In week 4, the strategy of 'clues in the text' was deleted and the list was renumbered. Week 4 results showed that once again instead of selecting 'other strategy', most test-takers selected another alternative from the modified checklist 'matching the alternative with the text' also indicating an instrument effect. In week 5 the checklist was not given and test-takers were requested to write a brief description of the strategy or strategies they used to answer each question. The data showed that test-takers used words and language that were different from the checklist. Allan (1995:151) concludes that the checklist "exercised to an unknown extent an instrument effect on the readers who used it, and that it biased the responses, introducing random error." However, there are two serious limitations to the study. The checklist was administered in English and the respondents unlike those in Nevo's study, were not trained to use it, therefore, test-takers with low ability in English may have not understood or may have misunderstood some of the descriptions in the checklist. Allan attempted to explain the results of the study by stating that "Chinese students are not used to being asked either to reflect on their strategies or to provide information about them" (p.150) and concluding that instrument effect may have a greater impact with some cultural groups than others. However, this is not a fair statement because there is no evidence to indicate that when given clear instructions, Chinese or even Egyptian learners cannot report on strategies used. Self-report instruments should be designed using

terms elicited from samples of the target population or in the respondents' native language. Moreover, rigorous verbal reporting methods should include sufficient guidance during data collection sessions as described by Green (1998).

In another study on reading tests, Anderson et al (1991) investigated the relationship between test-taking strategies, test tasks and performance on those tasks. The study involved 28 Spanish-speaking students enrolled in a university level ESL program. Students were required to take two forms of a standardized reading comprehension test which consisted of 15 reading passages and a total of 45 multiple-choice questions. The first test provided a measure of reading comprehension skills and during the second test, students provided retrospective think-aloud protocols after reading each passage and answering the items for that passage. The verbal report data were classified into 47 processing strategies which were grouped into the following categories:

- "1. Strategies for supervising strategy use (e.g. stating failure to understand a portion of text, confirming an inference)
2. Support strategies (e.g. skipping unknown words, skimming reading material for a general understanding)
3. Paraphrase strategies (e.g. translating a word or phrase into the L1, breaking lexical items into parts)
4. Strategies for establishing coherence in text (e.g. reading ahead, using background knowledge)
5. Test-taking strategies (e.g. looking for answers in chronological order in the passage, selecting an alternative through deductive reasoning)."

(Bachman & Cohen, 1998:17)

In addition to the verbal report data, a content analysis of the reading passages and questions was conducted based on the test designer's perspective and on a taxonomy available in the literature (Pearson & Johnson's 1978 Question and Answer relationships) and classical test item statistics: item difficulty and discrimination were obtained. The results showed that test-takers used some strategies consistently across the different types of items and also used other strategies differently depending on the type of question. The test performance data indicated that test-takers used less strategies on some of the easy items. Thus, the key finding is that strategy use is related to item ease or difficulty and the methodology used was retrospective think-aloud protocols.

Wijgh (1996) investigated the reading strategies of 13 Dutch secondary school students studying French using think alouds. The students were given a reading test that consisted of 13 multiple-choice questions on a variety of short authentic texts and were asked to think-aloud while they took the test. The researcher designed a protocol of reading strategies based on expected strategies to be used which was used to analyze the verbal protocols. The final list included a total of 17 strategies that were used by students. The researcher compared the students' actual reading strategies used with the "idealized, intended behavior as formulated by the researchers" (p.154). Most of the subjects used the same strategies of reading the question followed by reading the text globally and did not choose efficient reading strategies "corresponding to the reading objective intended by the test question" (p.161). This was explained by the researcher that the students had not been trained in developing reading strategies and that they may not have been efficient readers in the foreign language. Another possible conclusion is that these idealized strategies are not really used by test-takers in the actual test-taking situation. These strategies may be useful in learning or studying for a test.

Cloze reading tests

The two studies (Emanuel, 1982 and Hashkes & Koffman, 1982) reported by Cohen (1984) investigated the strategies used while taking cloze tests. One study (Emanuel, 1982) involved 25 Israeli ninth-grade EFL students and the other (Hashkes & Koffman, 1982) involved 22 Israeli twelfth-grade EFL students and four native English speakers. Retrospective verbal reports were used. In the two studies students were interviewed about how they responded to the cloze test immediately after taking the test. The results of the verbal report data indicated that only 25% of the test-takers reported reading the entire passage before answering, in spite of the fact that the instructions on the test very clearly requested students to initially read the whole text. The findings also indicated 16% of the test-takers reported not using the surrounding context of the preceding or following sentences for filling in the blanks. The majority of test-takers (64%) reported looking for clues to the answer within the same sentence. Most students reported using context 'part of the time' and only one-third reported using context 'all the time'. On investigating the strategies used by test-takers when faced with a blank they did not know how to fill, it was found that lower ability students left the space blank while higher ability students tended to guess an answer based on the immediate context. Another strategy reported was that of

translation. It was found that those who reported using translation while completing the task scored lower than those who did not. Therefore, there seems to be some differences in the strategies used by lower ability and higher ability test-takers and most test-takers used sentence level strategies.

The Hashkes & Koffman study also investigated the taking of cloze tests in a native language and they found that native speakers tended to use the context to find clues more than the non-native speakers. Native speakers also reported on extensive use of rereading sentences. These results differed from those of a study conducted by Kesar (1990) on 18 fifth-grade students at three levels of reading ability completing a cloze test in Hebrew L1. Students were asked to think aloud while completing the test and the verbal report protocols revealed at least 26 different strategies used. Although the results showed that the better readers tended to use the context more and did better on the task overall, all test-takers preferred using strategies at the sentence level. The strategies identified in this study "were grouped into seven categories: word level / part of sentence; sentence level; and five categories at the level of discourse-intersentential, whole-text level, extratextual level, metacognitive level, and 'other'" (Cohen, 1998: 105).

Cohen concludes from the strategy research in taking cloze tests that these tests "assess local-level reading more than they measure global reading ability" (p.105). Thus, two test-taking strategy studies in taking cloze tests yielded different results which means that more research using different methods and students of different is really needed in this area.

Summarization reading tests

Cohen (1994) investigated the strategies used by 5 native Portuguese speakers who had completed an EAP (English for Academic Purposes) course in a university in Brazil, in summarizing tasks. The test consisted of two parts. In Part I, test-takers were requested to choose two out of three short texts (400, 300 and 160 words respectively) to "indicate the topic treated by each one, in Portuguese" (p.177). Part II consisted of a text of 850 words and test-takers were required to indicate the topic and to identify the main idea in each section. The section is defined as "the text relating to each question in the article" (p.177). Test-takers were requested to think-aloud while taking the test and there was no time limit set. Test-takers were tested separately and a research assistant took notes and also wrote down all strategies observed. The time taken for the test ranged from 1.5 hours to 3 hours. The notes

from the verbal report protocols were analyzed and the strategies were identified using Sarig's (1987) taxonomy as follows:

1. Technical facilitation strategies including: underlining words, skipping material and dealing with difficult material by keeping the summary vague or including many details;
2. Clarification and simplification strategies including: translation, referring to the dictionary and interpreting idiomatic or technical words literally;
3. Coherence detection and production strategies including: using background knowledge, using explicit textual signals such as pronouns and conjunctions, using illustrations and using titles and subtitles;
4. Metacognitive monitoring strategies including: monitoring own performance and conscious planning.

The main problem identified was that of low reliability of raters and of the test. The two raters who scored the summaries differed in the assessment criteria they used and the instructions given in the test were not explicit. The instructions did not indicate how long the summaries should be, number of points for each or whether the summary should be in the form of a list or a paragraph. Thus, more research is required in the process of taking and rating summary tests.

Short answer reading tests

In a validation study of the Advanced English Reading Test of undergraduates in China, 69 Chinese students were required to think aloud onto tapes while taking the test (Weir et al, 2000). Test-takers were also asked to complete a checklist immediately after they finished the test. Students were to select the primary and secondary strategy (if any) used from a list of 5 strategies for each section of the test. The test consisted of 5 sections including 12 passages and 60 questions. Short answer formats were used with the amount of writing being strictly controlled. Answers were required to be no longer than 8 words. The item formats included short answers, table / flow-chart and sentence completion, writing summarizing sentences and banked cloze in section 5 where test-takers were required to fill in 5 blanks in each passage by choosing from a bank of 10 words. Results of the test-takers' retrospection data using the checklist showed that students found it difficult to distinguish between search reading and scanning. In general, there was low agreement between students' perceptions of strategies used in sections 1, 3 and 5 and those of the experts' judgment. This result is similar to that of Wijgh (1996) who

found that test-takers do not necessarily use the strategies intended by the test designers. On looking at the data from the top 21 students, their perceived strategies were more in agreement with the test developers' expectations. Thus, more proficient readers may have more awareness of strategy use. The verbal report protocols from 27 tapes were transcribed and analyzed. Based on their total scores, test-takers were divided into three ability groups: top, middle and bottom. The data indicated that the middle group most often used the expected strategies and in particular the expeditious reading sections. The top group processed a large amount of text in the time available. Some students in the bottom group were not aware of different strategies and just read carefully all 5 sections of the test. Other students in the bottom group, although aware of strategies, were unable to apply them because of low linguistic ability. Background knowledge did not seem to have an effect on test performance. Therefore, strategy use does seem to vary across different ability groups.

In this study, test-takers were also required to complete a questionnaire on their attitudes and reactions to the test. The questionnaire elicited data on test-takers perceptions of language difficulty of each passage, their familiarity with the topic, their interest in each passage, their familiarity with and attitudes towards the different formats, whether timing was sufficient, their perceptions of difficulty of each section, use of strategies in real life and prior training they had received in each strategy.

In all the test-taking strategy studies mentioned so far, lists of strategies used by test-takers were identified (Appendix 1 includes some of these lists) and in fact, Cohen (1998) presented a composite list of 18 strategies for taking a multiple-choice reading test based on the strategies mentioned in one or more of the studies he reviewed. Thus, the research has produced a series of strategies that test-takers use when taking multiple-choice reading tests. Several studies (Weir et al, 2000; Padron & Waxman, 1987; Gordon, 1987; Hashkes & Koffman, 1982) investigated the relationship between the specific use of test-taking strategies and performance on a language test (high and low ability learners) and in two other studies, strategies were classified as either contributing or not contributing to the correct answer (Nevo, 1989) or related to ease or difficulty of the item or text (Anderson et al, 1991; Nevo, 1989). The methods used to obtain data on test-taking strategies were either think aloud reports or checklists / questionnaires. In only one study useful data was obtained using both methods (Weir et al, 2000). Thus, a study that would combine both think aloud and a checklist / questionnaire would yield interesting results. Thus,

more research is needed on contributory and non-contributory test-taking strategies in relation to performance on a test using a combined approach of both think alouds and checklists / questionnaires. Furthermore, none of these studies have related these strategies to any processing or language testing model.

As mentioned above, besides the test-taking strategy research conducted in the area of reading, only one study was found in the area of listening comprehension and another study investigating strategies used in a language proficiency test.

2.2.3.2 Listening Tests

Buck (1994) investigated the processes in taking a language listening test utilizing retrospective verbal reports using a structured interview procedure. Six Japanese-speaking ESL students studying at a British university were asked to introspect while taking a short-answer listening comprehension test. The test consisted of 13 sections with 54 short answer questions. After each section, test-takers answered the questions and were then interviewed. Each interview was conducted in Japanese and lasted for two hours. The protocols were analyzed and the data showed that top-down processes are essential in listening comprehension. Test-takers differed from one another in their use of strategies and skills. Inferencing is used as a means to compensate linguistic deficiencies and a strong tendency for listeners to expectations which may or may not aid comprehension. All listeners were found to make mental images usually going beyond the text. Buck (1994) stated that "interest, emotional reaction and motivation can either facilitate or interfere with comprehension" (p.163).

2.2.3.3 Proficiency Tests

Purpura (1999) investigated the relationship between metacognitive and cognitive strategies and performance on language tests. The subjects were 1,382 test-takers in Spain, Turkey and the Czech Republic. The EFL proficiency test was the FCE (First Certificate in English) Anchor Test developed by UCLES (University of Cambridge Local Examinations Syndicate). This anchor test is used in equating different UCLES EFL tests and it is geared for intermediate level students. The test consisted of a reading comprehension section with 30 multiple-choice questions and a 'use of English' section with 40 limited-production items testing students' ability at the word and sentence levels. Section 2 included word formation, cloze and sentence

formation. The cognitive strategies questionnaire was designed based on a human information processing model of Gagne, Yekovich & Yekovich (1993 in Purpura, 1997) and on the work of researchers in the field (O'Malley & Chamot, 1990; Oxford, 1990). The metacognitive strategies questionnaire was designed based on Bachman & Palmer's (1996) model and on the work of O'Malley & Chamot (1990) and Wenden (1991). The responses to the test and the strategy questionnaires were analyzed using structural equation modeling. Purpura's study was the first study found where test-taking strategies were linked to a processing model and a questionnaire was used to collect the data.

The results showed that cognitive strategy use is a multi-dimensional construct consisting of comprehending, memory and retrieval processes. The study also showed that metacognitive strategy use was a unidimensional construct. Strategies such as "goal setting, planning, monitoring, self-evaluating and self-testing, often thought of as separate metacognitive strategies, all form part of one underlying construct involving assessment" (Purpura, 1999:178). The test data analysis yielded two factors of language ability: reading and a lexico-grammatical ability. The grammar ability correlated highly with reading ability which suggested that the reading test items invoked bottom-up reading processes. Metacognitive strategy use had no direct impact on test performance, however, it did have an indirect influence by means of cognitive processing i.e. using metacognitive strategies to invoke one or more of the retrieval processes did affect performance. Metacognitive strategy use had a moderate direct impact on comprehending processes and a strong direct effect on memory and retrieval processes. Metacognitive strategy use functions in an executive capacity in concert with cognitive processes. Cognitive processes had no effect on reading ability but did have a direct impact on lexico-grammatical ability. Comprehending had little effect on lexico-grammatical ability, memory had a large negative effect while retrieval had a strong positive impact. Thus, good test-taking strategies could mean "the ability to retrieve information from long-term memory without spending time trying to 'learn' or 'remember' during a test" (p.173). The results of the study also showed that metacognitive and cognitive strategy use differed between low ability and high ability groups. Low performers used a high degree of metacognitive processing in retrieving information. At lower levels of proficiency, more effort is required to process information indicating a lack of automaticity.

The findings indicated that strategies or clusters of strategies do have an impact on test performance and thus, test-takers "need to know how to use both metacognitive and cognitive strategies effectively" (p.180). Purpura also proposed a test-taking style construct which he described as "a continuum ranging from product-oriented to process-oriented, where product-oriented test-takers are able to answer questions quickly and efficiently by retrieving information from long-term memory, while process-oriented test-takers might be more prone to spending time trying to comprehend or remember test input, rather than simply answering the question being asked" (p.181).

2.2.4 Summary of Test-taking Strategy Studies Reviewed

Table (2) shows a summary of all the language test-taking strategy studies reviewed above. Allan's study was not included because of the limitations identified in the study.

Table (2): Summary of test-taking strategy research studies

Study	Subjects	Strategy Data Collection Methods	Findings
Multiple-choice reading tests			
Larson (1981)	40 Israeli ESL university students	Retrospective interviews	<ul style="list-style-type: none"> • Three strategies identified. • Overall approach is to read questions first or just read parts of passage corresponding to the questions.
Israel (1982)	57 Israeili EFL students	Performance on test	<ul style="list-style-type: none"> • Students rely on background knowledge rather than read passage carefully
Gordon (1987)	30 tenth-grade Israeli students	Think-aloud protocols	<ul style="list-style-type: none"> • Identified global level & local (sentence/level) strategies • High ability and low ability students used different strategies
Waxman & Padron (1987)	82 Grades 3,4, & 5 Hispanic ESL students	Retrospective questionnaire	<ul style="list-style-type: none"> • Low ability students use inappropriate cognitive reading strategies

Table (2): Summary of test-taking strategy research studies (continued)

Study	Subjects	Strategy Data Collection Methods	Findings
Nevo (1989)	42 tenth-grade Israeli students studying French	Response strategy checklist	<ul style="list-style-type: none"> • 8 contributory and 7 non-contributory strategies were identified • Transfer of strategies from L1 to L2 • Use of more non-contributory strategies in L2 • More use of contributory strategies in easy texts
Anderson et al (1991)	28 Spanish-speaking ESL university students	Retrospective think-aloud protocols (taped)	<ul style="list-style-type: none"> • 47 strategies identified • Some strategies used differently across different question types • Less strategies used on some easy items
Wijgh (1996)	13 Dutch secondary school students studying French	Think-aloud protocols (taped)	<ul style="list-style-type: none"> • 17 strategies identified • Most students read the questions, then read text globally and did not use efficient strategies
Cloze reading tests			
Emanuel (1982)	25 ninth-grade Israeli EFL students	Retrospective interviews	<ul style="list-style-type: none"> • Students do not read entire passage before answering
Hashkes & Koffman (1982)	22 twelfth-grade Israeli EFL students & 4 native English speakers	Retrospective interviews	<ul style="list-style-type: none"> • Different strategies used by low ability and high ability students • Use of sentence level rather than discourse-level strategies
Kesar (1990)	18 fifth-grade native speakers of Hebrew	Think-aloud protocols (taped)	<ul style="list-style-type: none"> • 26 strategies used • Better readers used context more • Most test-takers preferred sentence level strategies
Summarization reading tests			
Cohen (1994)	5 Brazilian students who completed an EAP course	Think-aloud protocols (not tape-recorded) – took extensive notes	<ul style="list-style-type: none"> • Identified strategies used based on a taxonomy of strategies • Problems with reliability of raters and test

Table (2): Summary of test-taking strategy research studies (continued)

Study	Subjects	Strategy Data Collection Methods	Findings
Short answer reading tests			
Weir et al (2000)	69 Chinese undergraduate students	Think-aloud protocols (27 tapes analyzed) Retrospective checklist	<ul style="list-style-type: none"> • Top group readers process text at very fast speed • Most middle group readers use strategies as expected by test developers • Bottom group readers are unaware of strategies or do not have linguistic ability to use them. • Background knowledge had no effect on test performance,
Listening tests			
Buck (1994)	6 Japanese ESL students studying at a British university	Immediate retrospective structured interviews (taped)	<ul style="list-style-type: none"> • Successful listening involved a variety of subskills which differed from one test-taker to another • Top-down processes are crucial to listening comprehension • Inferencing is an integral part of listening
Proficiency tests			
Purpura (1999)	1,382 EFL test-takers in Spain, Turkey & Czech Republic	Retrospective metacognitive and cognitive questionnaires	<ul style="list-style-type: none"> • Cognitive strategy use is a multi-dimensional construct • Metacognitive strategy use is a unidimensional construct • Metacognitive strategy use exerts an executive function on cognitive strategy use • Both metacognitive and cognitive strategy use impact test performance • Strategy use differs between low ability and high ability test-takers • A test-taking style construct is proposed

In conclusion, most of the test-taking strategy research has resulted in lists of strategies that are used by test-takers when taking different types of reading tests and an open-ended listening test. Several studies attempted to link strategy use with performance on a test and the findings of these studies (Gordon, 1987; Hashkes & Koffman, 1982; Weir, 2000; Purpura, 1999) showed that high ability and low ability test-takers use different strategies. However, only one study (Purpura, 1999)

approached the issue of strategy use and language test performance from a model-oriented perspective. None of the other studies attempted to relate strategy use to any information processing, cognitive or language assessment model. None of the studies investigated the strategies used during a writing test. Furthermore, only one study (Weir, 2000) used two types of strategy data collection methods combining think-aloud protocols and a retrospective checklist which was designed based on the strategies that the test designers expected the test-takers to use. Only two of these test-taking strategy studies (Weir, 2000 and Buck, 1994) included test-takers' reactions or attitudes while taking tests. Therefore, more research is needed in investigating test-taking strategies and test-takers' reactions or attitudes while completing different test tasks: writing, reading, listening, summarization, cloze,... using both think alouds and checklists / questionnaires. Furthermore, the findings should be related to or interpreted within a language testing model.

One of the implications of Purpura's work is that the metacognitive component of Bachman & Palmer's (1996) model needs to be reconceptualized and thus, the metacognitive component consists of one element: assessment. This is described by Purpura (1999: 178) as follows:

"... when a learner sets goals, she *assesses* what she wants to achieve; when she plans what to do next, she *assesses* the situation and *assesses* which actions to pursue; when she monitors her work, she *assesses* how she is carrying out the task at hand; when she evaluates her work, she *assesses* the quality of her actions; and when she tests her knowledge or understanding of something, she *assesses* what she thinks she knows or understands."

This study builds on the work of strategy researchers by including strategy use within a modified Bachman & Palmer (1996) model. Both think alouds and a questionnaire will be used to investigate both test-taking strategies and affective factors while students are taking a proficiency test that includes listening, reading and writing tasks. The construct of the proficiency test will be empirically defined to match the context of the study (i.e. Egyptian adult learners studying English at the Center for Adult & Continuing Education of the American University in Cairo) and the findings of the test-taking strategies and affective factors will be interpreted within a language testing model.

2.3 AFFECTIVE FACTORS AND TEST PERFORMANCE

In language testing, research in affective factors and test performance has also been limited in scope and in all the studies previously mentioned (Cohen, 1984; Zeidner & Benoussan, 1988; Bradshaw, 1990; Brown, 1993; Peirce & Stein, 1995; Weir et al, 2000) the test-takers' reactions to specific language tests or test item types have been investigated with the aim of obtaining feedback to be used in further refining or developing the tests or test items themselves. None of these studies investigated affective factors and test performance within a process approach or within a language assessment model.

Roizen's (1982, in Cohen; 1984) study showed that 60% of the students did not like the cloze test. Zeidner & Benoussan (1988) found that students prefer written English text comprehension tests to oral tests and that females experience more anxiety in taking oral language tests than males. However, Bradshaw (1990) investigating students' reactions to a grammar, reading comprehension and C-test found no significant gender differences and that students reacted negatively to the C-test. Brown (1993) explored test-takers' reactions regarding difficulty, adequacy of time given for preparation and response and general comments to a tape-mediated oral test of Japanese. The results showed no gender differences in reactions to the oral test and that higher ability students perceived the test to be less difficult. Pierce & Stein (1995) investigated the suitability of a reading comprehension passage about monkeys in an English proficiency test for black high school students. The study revealed negative reactions and that most students viewed the passage as racist, however, 63% of the students scored 80% or more. Results of Weir et al's (2000) study showed that in general, the higher the scores: the less difficult the passage, the more familiar they are with the topic, the more interesting the text is perceived, the more sufficient the time is seen, the more frequently they used the strategies tested in real life and the more training they had received in a particular strategy. Therefore, only one study (Weir et al, 2000) showed a clear relationship between perceptions or affective factors and performance on a language test.

Another area of research that focuses on adverse reactions of test-takers while taking language tests is that of anxiety research. A considerable amount of research has been conducted on language anxiety in language learning in the classroom (for example, Horwitz & Young, 1991; Aida, 1994; MacIntyre et al,

1997), on the effect of language anxiety on test performance (Horwitz & Young, 1991; Phillips, 1992) and on the relationship between anxiety and language proficiency (Clement & Kruidenier, 1985; Gardner et al, 1989; MacIntyre, 1995 and MacIntyre et al, 1997). The research has shown that excessive anxiety does have a negative effect on test performance and that "the interaction between personality, achievement, anxiety and the nature of tests and the testing situation is very complex and as yet only partially understood" (Bradshaw, 1990:15). Wine (1971, in Bradshaw, 1990) after having surveyed the research done at that time found that students "who are most anxious about their performance are likely to be diverted from task-relevant activities, and that this diversion will disrupt information processing" (p.15). Madsen (1982, in Bradshaw, 1990) found debilitating effects of anxiety on the test performance of adult EFL students. The research has also shown that language anxiety correlates negatively with proficiency and that as proficiency increases, the level of anxiety tends to decrease.

Although several empirical studies have demonstrated a relationship between language anxiety and achievement, however, some researchers have had difficulty in describing the role of anxiety in second language learning. In the research there is a distinction between debilitating and facilitating anxiety. Debilitating anxiety impairs performance while facilitating anxiety "mobilizes resources to accomplish a task" (Ehrman, 1996:184). The Yerkes-Dodson Law describes a curvilinear relationship between anxiety and task performance. This relation is represented by an inverted U on a graph with anxiety on the horizontal axis and performance on the vertical axis. The two extremes represent too much or too little anxiety, with a peak representing optimum performance at some level of anxiety. MacIntyre (1995) describes a model where there is a cyclical or recursive relationship between anxiety, cognition and behavior. To illustrate the model, the following example is described in a language classroom: a student is asked to answer a question in class. This may cause the student to become anxious and his attention becomes divided between the task and his reactions to it. The result is that "cognitive performance is diminished because of the divided attention and therefore, performance suffers, leading to negative self-evaluations and more self-deprecating cognition which further impairs performance, and so on" (p.92).

Therefore, having identified the gaps in the test-taking strategy literature to date, this study investigates test-taking strategies including affective strategies within

Bachman & Palmer's conceptual framework combining qualitative verbal protocol data and quantitative questionnaire data. The research questions that emerge from the literature review and specifically address these gaps are:

Research Question 2:

- What are the cognitive and metacognitive test-taking strategies used by Egyptian adult EFL learners in a specific context when taking a placement test?

Research Question 3:

- What is the relationship between affective factors, test-taking strategies and test performance? Further questions that focus on the relationship between test-taking strategies, affective factors and test performance are:

Do test-taking strategies vary across different levels of test performance?

Do affective factors vary across different level of test performance?

How do test-taking strategies differ across different levels of test performance?

How do affective factors differ across different levels of test performance?

Since self-assessment is an important component in the Bachman & Palmer model and in order to posit a modified model that incorporates test-taking strategies including affective strategies, the following section describes the considerable amount of research conducted in the area of self-assessment.

2.4 RESEARCH IN LANGUAGE SELF-ASSESSMENT

One of the first people to mention the use of self-assessment in measuring language ability was Upshur in 1975. He stated that learners can tap into their whole range of language skills, whereas language tests can only sample a limited range. This interest was in line with the development of the student-centered and self-directed approaches in language learning. Self-assessment of language proficiency is concerned with learners assessing their own ability in the language. The capacity for self-assessment is a key element in any area of competence and autonomy in learning. A basic component of language competence is the “metacognitive ability to assess linguistic needs, resources and communicative success” (Myles, 1997).

In spite of the theoretical support for the importance of self-assessment of language proficiency, a review of the research done in this area revealed

contradictory results with the conclusion that “no consensus has been reached on the merits of the self-assessment approach” (Oscarson, 1997: 182).

2.4.1 Validation Studies Supporting Self-Assessment

Most of the studies have focused on the validity and reliability of self-assessment and self-assessment instruments. Many of them involved adult learners who were studying English, French or Dutch as a second or foreign language. In 1989, Blanche & Merino conducted an extensive review of the literature of self-evaluation of foreign language skills. They reviewed sixteen (16) studies. They concluded that there was consistent overall agreement between self-assessment and other external measures. They also found that accuracy of self-assessment varied but in general, ranged from good to very good. Coombe (1992) found a strong relationship ($r=.83$) between self-assessment ratings and functional literacy skills of Russian, Vietnamese and Cambodian refugees in the United States. Smith & Baldauf (1982) also found a strong relationship between self rating and trained interviewer ratings for migrants to Australia.

Ross (1998) conducted a meta-analysis of ten studies on self-assessment and found that the number of subjects in the studies varied considerably. He concluded that self-assessment is quite robust and the results reported were:

“self-assessment of receptive skills is more accurate than that of productive skills; self-assessment of achievement (based on experience ‘can do’) is more accurate than self-assessment of proficiency (‘could do’); instructor assessments of student proficiency are more accurate than self-assessment, but not by much; and self-assessment offers a practical alternative to formal assessment for ‘low stakes’ testing needs.” (p.3)

2.4.2 Validation Studies not Supporting Self-Assessment

However, there have been several studies that have reported small or no significant relationships between self-assessment and other external measures of language ability. Janssen-van Dielen (1992) conducted a study on learners of Dutch as a second language. A C-test used as an anchor test, a self-assessment test and a questionnaire eliciting information on background variables were administered to 973 testees. The results showed that there was no consistent significant relationship between performance on the two tests. There were no strong relationships between

background variables, age, gender, country of origin, length of residence, education level and the amount of language training as indicated by attendance of language courses, and self-assessment. In fact, “disquieting are the rather low but in most cases negative correlations between the amount of language training and correct estimation (p.42)”. Blanche & Merino (1989) reported on two studies conducted by Anderson (1982) and Blanche (1985) that showed that there no significant relationships between the accuracy of self-assessments of language proficiency and their actual classroom/test performance. Peirce et al (1993:26) mentioned a study conducted at the University of Ottawa by Ready-Morfitt (1991) which concluded that self-assessment as a placement tool “can be unreliable if there is a clearly perceived advantage in mis-assessment” and that “self-assessment no longer enjoys such a high rate of success at the university.” Heilenman (1991) reported that studies conducted by Blue (1988), Oller & Perkins (1978) and Wesche et al (1990) also showed small or no statistically significant relationships between self-assessment and other proficiency measures.

The results of self-assessment research are not conclusive. The uses of self-assessment for language learners have been varied, ranging from self-assessment used as a placement test (LeBlanc & Painchaud, 1985) to self-assessment for formative purposes. Bachman & Palmer (1989) state that self-assessment studies differ in design and purpose. They compared their findings with the results of Davidson & Henning’s (1985) study on the applicability of Rasch modelling to the development of rating scales. Davidson and Henning concluded that little confidence should be placed on self-assessments while Bachman & Palmer (1989:22) found that “self-ratings can be reliable and valid measures of communicative language abilities.” Davidson & Henning used a different measure that was not tested and students were asked to rate their own difficulty with 11 English language skills using a seven-point rating scale : none, very little, some, average, more than average, much, extreme. The questions used were all of the same type. Another difference was that Davidson & Henning used IRT modelling which is more suited to test “the extent to which individual items fit a single underlying dimension” (p.23) while Bachman & Palmer used a multitrait-multimethod design and confirmatory factor analysis.

2.4.3 Accuracy of Self-Assessment

There has been concern about the ability of learners to accurately assess their own language ability. Several studies have reported a tendency of less proficient learners to overestimate their abilities (Davidson & Henning, 1985; Janssen-van Dielen, 1992) and that more proficient learners underestimated their linguistic abilities (Blanche & Merino, 1989). Blanche (1990) conducted a study investigating the accuracy of self-assessments using a regular test as the primary self-evaluation tool and focusing on only one skill, speaking. There were two groups of students : 11 advanced students and 32 students at the beginner stage. He also investigated the impact of language experience and language learning strategies and attitudes on self-assessment. The results matched previous studies where weak students overrated themselves, the overall self-assessment were good, however, high achievers were extremely accurate, in contrast to findings of other studies. There was no significant relationship between learners' self-assessment of their language proficiency and their actual performance in the classroom (achievement). There was no relationship between the variables: language experience, strategies and attitudes, and self-assessment. Blanche stated that this study had many limitations in terms of the subjects, procedures and setting. Bachman & Palmer (1981) and Blanche (1990) stated that many students found it difficult to evaluate their grammatical ability. In another study by Heilenman (1991) investigating the role of response effects in self-assessment, the results showed that acquiescence and overestimation were present. The subjects were 232 students of French at the University of Iowa who were administered a 65-item self-assessment questionnaire. The questionnaire consisted of four scales : grammar, vocabulary, accuracy and fluency. Students were to respond to the questions using a 5-point response scale: strongly agree, agree, neither agree nor disagree, disagree and strongly disagree. Two types of questions were used: can-do and difficulty. The results showed that "overestimation as measured in relationship to instructor judgment, is a real phenomenon" (p.188). Overestimation was most evident for less experienced learners. Furthermore, less experienced students are more liable to acquiescence (a tendency to respond positively) than more experienced students.

2.4.4 Design of Self-Assessment Instruments and Benchmarking

Another issue that has been extensively debated in the self-assessment literature is related to the design of the instruments and the use of benchmarks. Oscarson (1978:14) stated that most learners will be able to assess their own language ability, “provided that they have at their disposal a measuring standard by which they can express their intuitions.” LeBlanc & Painchaud’s (1985) study showed that the self-assessment results improved (as measured by the correlation between self-assessment and a proficiency test score) when the questionnaire items used situations closely related to students as potential second language users. For example, the situations of reading posters on campus or information in a university calendar were better than general understanding of texts on the questionnaire. Bachman & Palmer (1989) found that the most effective question type asked about perceived “difficulty with production”. The least effective question type was the “can-do” question. In Heilenman’s study, no difference was found between ‘can-do’ and ‘difficulty’ questions.

Peirce et al (1993) conducted a study on 500 Grade 8 students in two different French immersion programs in Canada. The instruments used were a questionnaire and French proficiency tests. The questionnaire included two benchmarks: perceived language proficiency of francophone peers and difficulty with specific tasks. The results showed that there was a weak correlation between self-assessment and tested proficiency. On comparing effectiveness of benchmarks, it was found that situational benchmarks produced higher correlations than a more global benchmark. This result is in agreement with both LeBlanc & Painchaud’s and Bachman & Palmer’s studies. Another result was that using the ‘francophone peer’ benchmark students’ rated themselves more proficient in receptive skills (reading and listening) than productive skills (speaking and writing). However, with the ‘specific tasks’ benchmark students rated themselves as more proficient in literacy tasks (reading and writing) than oral tasks (listening and speaking).

2.4.5 Process of Self-Assessment

Moritz (1995) explored the cognitive processes and social-situational influences underlying students’ self-assessment of their language proficiency. The subjects were 28 learners of French at different course levels. A self-assessment questionnaire on the four language skills was used and verbal report data was

obtained using a think-aloud protocol and interviews. The results showed that six factors influenced self-assessment as follows: question interpretation, language learning background/experience, reference points, questionnaire-completion strategies, level of certainty about answers and level of self-confidence. It was also found that students used a variety of reference points or benchmarks when evaluating their own language abilities: social category, meaningful other, autobiographical and social context. These benchmarks are defined as follows:

“Social category is the standard as defined by the average performance of subjects’ current or past language learning colleagues. ... Meaningful other is a standard defined by the performance of an individual who is meaningful to the subject.... Autobiographical is a standard defined by the subject’s own past performance” and “social context is defined by the performance of the immediate context of people to whom the subject is currently exposed.” (p.1)

Self-assessments were influenced by individuals’ experiences. Moritz concludes that it is not possible to match individuality to a rating scale and that self-assessment is not useful as a placement tool because of compromises to validity.

The studies reviewed so far have mainly focused on the use of self-assessment as a means of indirectly assessing foreign language ability. The results are contradictory with some studies showing the correlation between self-assessment and other proficiency tests to be satisfactory (i.e. LeBlanc & Painchaud (1985) found a correlation of 0.53; Janssen-van Dieten (1989) reported correlations between 0.60 and 0.79; Coombe (1992) found a correlation of 0.83) while other studies have reported low correlations (Peirce et al (1993); Janssen-van Dieten (1989) and Wesche (1993) reported weak correlations between self-assessment and proficiency tests). Assuming that the self-evaluation instrument is sufficiently reliable, this suggests that the instrument is not valid to measure the same underlying trait as that of the proficiency test implying that different traits are being considered here. Most of the studies used a quantitative methodology and relatively little attention has been given to studying the process of self-assessment and feelings or affective factors were not investigated explicitly.

As previously mentioned, only three studies (Moritz, 1995; Peirce et al, 1993; Heilenman, 1991) have investigated the process of self-assessment and no models were proposed by the researchers. Moreover, these three studies were conducted on

learners of French in the U.S. and in Canada. No studies were found on the process of self-assessment in an EFL context that could be relevant to the social / cultural context in Egypt. Moritz's (1995) research which focused on identifying the factors influencing self-assessment using a qualitative methodology showed that these factors are "numerous, complex and interrelated". She concluded that "though there may be a few detectable patterns in the data, the overall impression is one of variation between and within individual subjects' responses". Moritz did not provide or define a conceptual framework for language self-assessment. Heilenman (1991) attempted to explain the inconsistent results and response effects in self-assessment research by referring to the Information Processing Model of cognitive tasks. However, as Moritz (1995) noted the model itself cannot explain the presence of response effects and it is not a valid representation of the cognitive processes involved in self-assessment. Therefore, it is concluded that the language self-assessment literature does not contribute to the further development of a modified Bachman & Palmer (1996) model.

The purpose of this thesis is to expand the Bachman & Palmer language testing model to include the interplay between test-taking strategies and affective factors while taking a test. Since the language assessment literature did not yield sufficient information to posit a modified model incorporating test-taking strategies and affective strategies, the researcher decided to extend the literature review beyond language assessment. Thus, a review was conducted of the relevant literature in educational psychology, specifically in the area of self-regulation. The following section critiques the literature focusing on self-regulation while taking tests, in order to shed light on the relationship between affective factors and test performance with the purpose of proposing a language testing processing model that is based on the modification of the Bachman & Palmer model.

2.5 SELF-REGULATION WHILE TAKING TESTS

2.5.1 Definition of Terms in Self-Regulation

Self-regulation is a relatively new area in psychological research and most of the research was conducted in the area of social psychology and personality in the 1980s. In the 1990s self-regulation constructs began to be applied in various fields such as education. Research in self-regulation has proliferated in the last few years and many different perspectives on self-regulation have been presented. For

example, researchers differ in their definition of self-regulation and use slightly different terminologies. In order to draw upon concepts in self-regulation that could be applied in this study it was important to identify a source in the literature that had reviewed the current state of the field and had arrived at some consensus regarding the conceptualization of the key components of self-regulation. The Handbook of Self-Regulation published in 2000 provided such an overview that was very much needed in order to provide guidance in defining the different constructs referred to in this study. In defining self-regulation, according to Zeidner, Boekaerts & Pintrich (2000:751) there is some consensus that “self-regulation involves cognitive, affective, motivational and behavioral components that provide the individual with the capacity to adjust his or her actions and goals to achieve desired results in light of changing environmental conditions.” Self-regulation involves a feedback loop and is cyclical because feedback obtained from one performance enables the individual to adjust their current behavior. The distinction between self-regulation and metacognition is not clear and there is considerable ambiguity about the relationship between these two terms. In order to clarify this overlap, the following definition is adopted in this study:

Metacognition is generally defined as the awareness individuals have of their personal resources in relation to the demands of particular tasks, along with the knowledge they possess of how to regulate their engagement in tasks to optimize goal-related processes and outcomes...Self-regulation may be viewed as the more comprehensive term, embracing both metacognitive knowledge and skills, as well as motivational, emotional, and behavioral monitoring and control processes.

Zeidner, Boekaerts & Pintrich (2000:752)

There is also considerable overlap in the definitions of self-regulation and coping. Both concepts involve processes in attaining personal goals. For the purposes of this thesis while self-regulation is the more comprehensive term, the research questions focus on metacognitive, cognitive (test-taking) and affective or emotional regulation strategies separately. Coping is defined as the “appraisal of the potential threat a situation poses to the person, its related emotional reactions, and the various procedures, mental actions and overt actions taken to manage the problem and the feelings it evokes.” (p.752). Researchers also disagree about the role of self-efficacy and affect in self-regulation. Zimmerman (2000), Schunk &

Ertmer (2000) view self-efficacy as part and parcel of self-regulation while researchers such as Endler & Kocovski (2000) do not view self-efficacy as a component of self-regulation. Self-efficacy is defined as the beliefs an individual has about one's capabilities in performing or attaining a personal goal. Researchers such as Pintrich (2000), Zimmerman (2000) and Carver & Scheier (2000) view affect or emotions as an integral part of the self-regulatory process while Vancouver (2000) does not. In this study both self-efficacy and affect or emotions are viewed as integral components of self-regulation. The important roles of self-efficacy and affect will clearly be seen in the following discussion about two key self-regulation models. Furthermore, the terms emotions and affect are used interchangeably in this thesis. Oatley & Jenkins (1996:124) state:

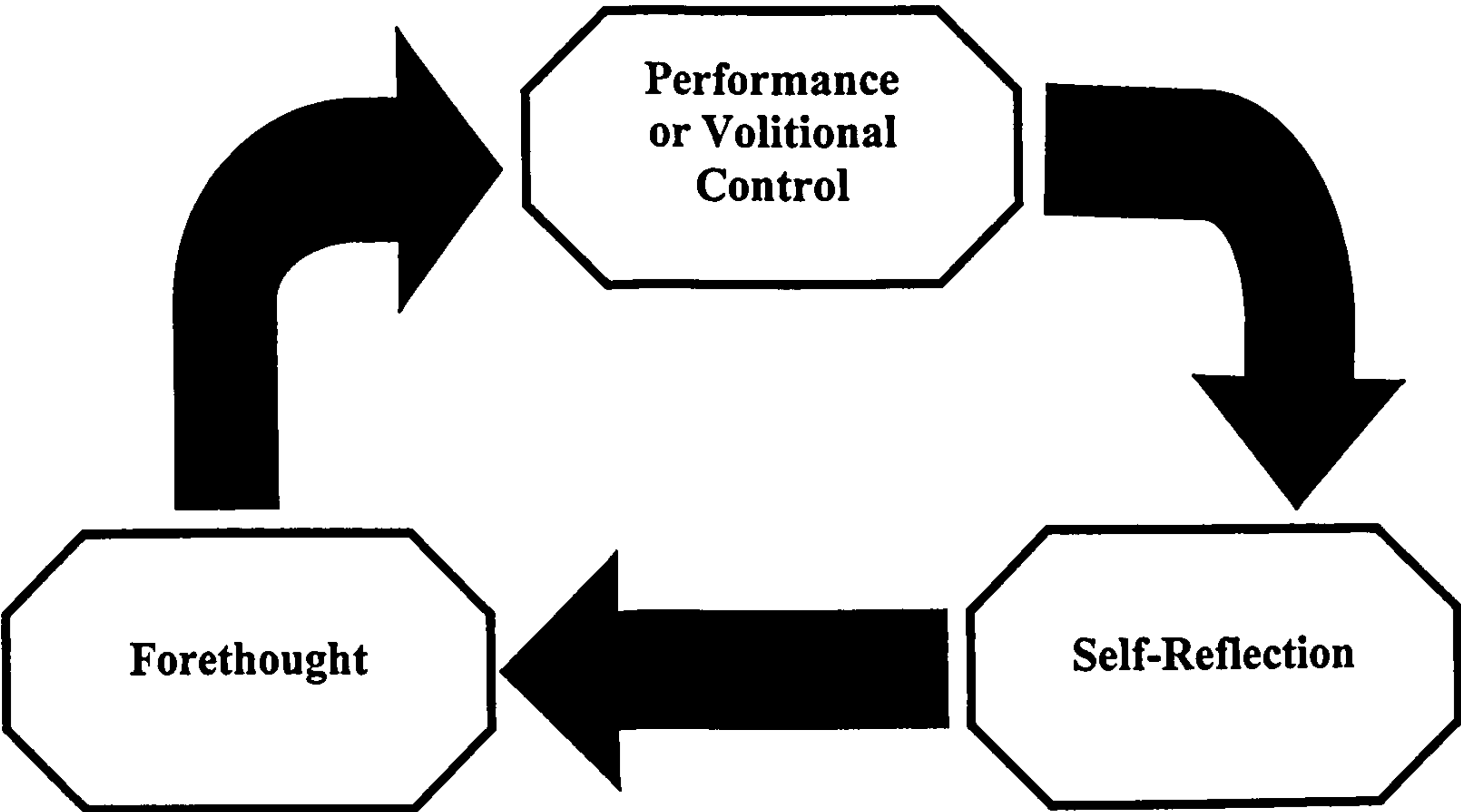
“The term ‘feeling’ is a synonym for emotion, although with a broader range. In the older psychological literature the term ‘affect’ was used. It is still used to imply an even wider range of phenomena that have anything to do with emotions, moods, dispositions, and preferences.”

2.5.2 Self-Regulation Models

In the literature, the two most prevalent models of self-regulation are the social-cognitive perspective and the control perspective. In the social-cognitive model, self-regulation consists of three cyclical phases: (1) forethought, (2) performance or volitional control and (3) self-reflection. The model is shown in Figure (3).

Figure (3): The social-cognitive model of self-regulation

(Zimmerman, 2000: 16)



Each of these phases is further broken down into sub-processes as shown in Table (3) taken from Zimmerman (2000:16).

Table (3): The social-cognitive model of self-regulation

CYCLICAL PHASES OF SELF-REGULATION		
FORETHOUGHT	PERFORMANCE OR VOLITIONAL CONTROL	SELF-REFLECTION
Task analysis <ul style="list-style-type: none">• Goal setting• Strategic planning Self-motivation beliefs <ul style="list-style-type: none">• Self-efficacy• Outcome expectations• Intrinsic interest / value• Goal orientation	Self-control <ul style="list-style-type: none">• Self-instruction• Imagery• Attention focusing• Task strategies Self-observation <ul style="list-style-type: none">• Self-recording• Self-experimentation	Self-judgment <ul style="list-style-type: none">• Self-evaluation• Causal attribution Self-reaction <ul style="list-style-type: none">• Self-satisfaction / affect• Adaptive-defensive

Phase (1): forethought consists of two main sub-processes: task analysis and self-motivation beliefs. Task analysis involves setting goals and strategic planning. Setting goals means deciding on specific learning or performance outcomes while strategic planning involves selecting planning and selecting strategies appropriate for the task. Planning and selecting strategies must be continually adjusted in view of diverse and changing personal, behavioral and contextual conditions. Underlying setting goals and strategic planning are a set of four self-motivation beliefs: self-efficacy, outcome expectations, intrinsic interest / value and goal orientation. Self-efficacy is the belief people have about their ability to perform or learn effectively and outcome expectations are the beliefs people have about the end results of learning and performance. For example, if a person has high self-efficacy and believes that she can obtain a score of A on a course and she has high outcome expectations that this score will have positive consequences such as a promotion on the job, then she will exert and sustain self-regulatory efforts to attain this goal. Thus, “the more capable people believe themselves to be, the higher the goals they set for themselves and the more firmly committed they remain to those goals” (Zimmerman, 2000:18). In this way a goal orientation results in intrinsic motivation.

Phase (2): performance of volitional control consists of two main sub-processes: self-control and self-observation. Self-control involves self-instruction, imagery, attention focusing and task strategies. To assist in performance, self-instruction is verbalizing how to complete a task and imagery involves creating mental pictures. Attention focusing means improving concentration and eliminating diversions and task strategies include study and performance strategies such as note taking and problem solving designed to assist performance. Self-observation includes self-recording and self-experimentation which “refer to a person’s tracking of specific aspects of their own performance, the conditions that surround it, and the effects that it produces” (Zimmerman, 2000:19).

The self-reflection phase (3) consists of self-judgment and self-reactions. Self-judgment involves self-evaluation and causal attribution. Self-evaluation refers to comparing one’s performance with a standard or with previous performance. Causal attribution involves judging whether poor performance is due to lack of ability or lack of effort. Attributions also depend on cognitive appraisal of the environmental circumstances or conditions. For example, when test-takers receive

low scores on a test, those who have high self-efficacy would more likely attribute this to lack of studying while test-takers who felt that the test was administered under adverse conditions, may attribute this to bad luck rather than lack of ability. Forethought may also affect attributional judgments. Test-takers who plan to use a specific strategy and then actually use it may attribute poor performance to that strategy rather than lack of ability. Self-reactions involve self-satisfaction and adaptive or defensive inferences. Self-satisfaction means satisfaction / positive affect and dissatisfaction / negative affect means anxiety with one's performance. The level of self-satisfaction depends on the value placed on the task and the greater the value the higher the self-satisfaction. "Adaptive or defensive inferences are conclusions about how one needs to alter his or her self-regulatory approach during subsequent efforts to learn or perform" (p.23). Adaptive self-reactions include changing goals or selecting more effective strategies and defensive self-reactions involve helplessness, avoiding the task and procrastination. Self-reactions in turn affect forethought processes reflecting the cyclical aspect of the model.

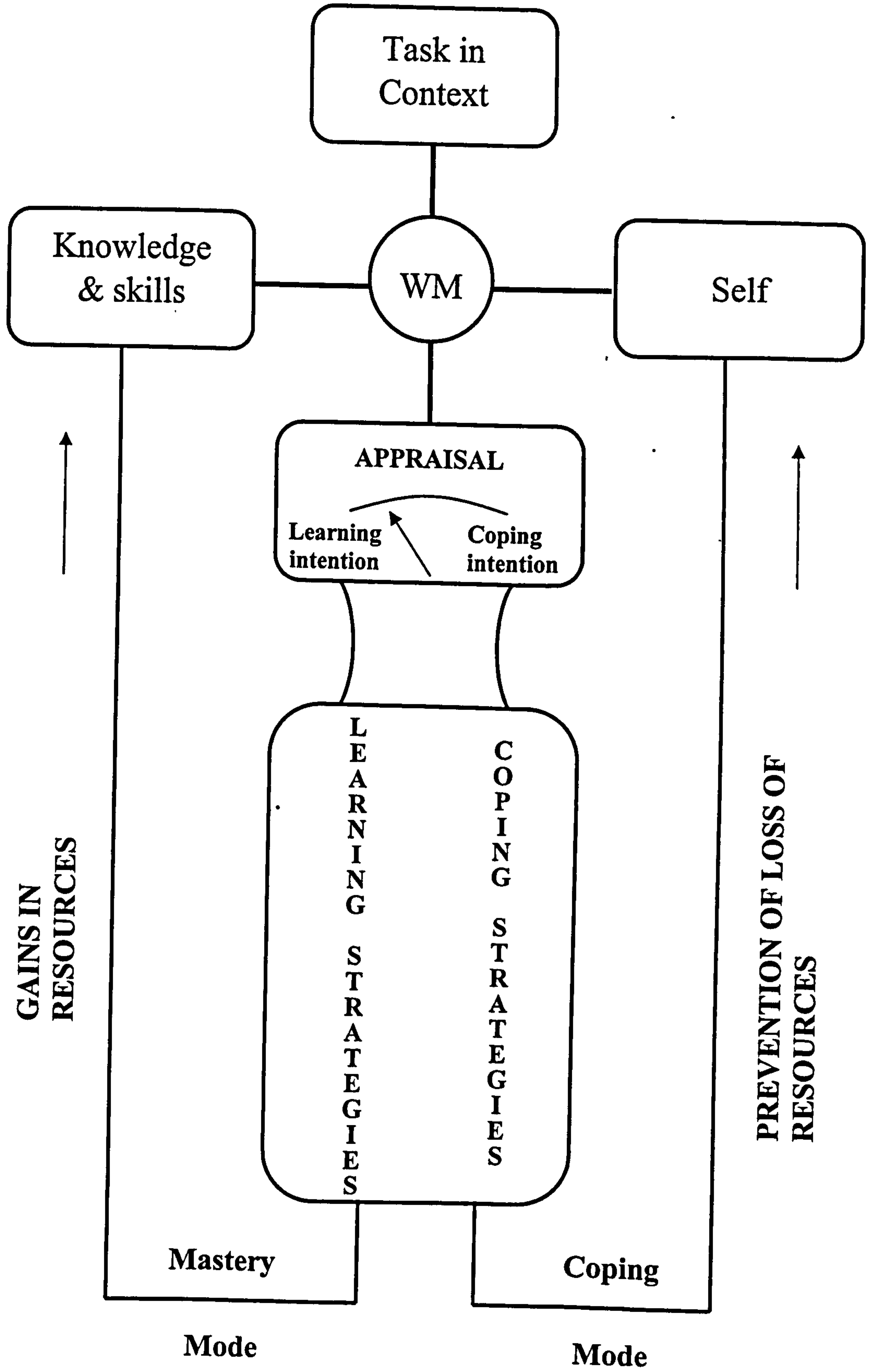
Within this model, the social milieu and the environment are viewed as sources for enhancing forethought performance or volitional control and self-reflection. For example when self-evaluating, the standard used to judge one's performance is usually formed from the social milieu while self-reactions can be enhanced by using environmental supports such as rewarding or praising oneself with breaks or other self-administered rewards. In language learning / testing another source used in self-evaluation also related to the environment and possibly culture is a person's perceptions of what it is to know a language i.e. how people in a specific context define the construct of language proficiency. Dysfunctions in self-regulation are explained in the model by the use of ineffective forethought and performance control techniques. For example, poor self-regulators are not able to plan strategically and thus, "they try to correct themselves using post hoc task outcomes" which are usually too late. These unfavorable outcomes results in self-dissatisfaction and defensive self-reactions in turn leading to low self-efficacy about future performance.

The second key model of self-regulation referred to as the Model of Adaptable Learning (Boekaerts & Niemivirta, 2000:429) is shown in Figure (4). It breaks down self-regulation into different control processes: metacognitive control, motivation control, emotion control and action control. It describes the continuous

process of self-assessment or appraisal during any learning process. These non-stop assessment processes and the emotions and strategies they elicit are part of a dynamic process. The internal working model (WM) of reality is the frame of reference of an individual in any learning situation. This WM is based on the three main sources: task demands, knowledge and skills relevant to the task and the learner's self. The learner must analyze the task and infer its difficulty and the skills required for its successful completion. The knowledge and skills include cognitive strategies that have been used successfully and metacognitive knowledge relevant to the task. The third source is related to the learners' self-system and includes their goals, and motivational beliefs activated by the situation. This model differentiates between a person's metacognition and an individual's motivational beliefs i.e. distinguishing between metacognitive and motivational control. Based on the WM, the learner appraises the learning situation and for every learning context, the appraisal is unique because the information obtained from the three sources may differ each time. These appraisals are continuous and involve both positive and negative feelings and cognitions. The result of the appraisal process may or may not lead to a discrepancy between the task demands, the knowledge and skills and self. If there is no discrepancy, no intense positive or negative emotions are experienced during the task. If there is a discrepancy, it may result in an increase in positive emotions where the learner views the situation as a challenge and leads him onto the mastery route (gaining resources). A discrepancy may also result in negative emotions or distress which leads the learner onto the coping route to restore well-being (preventing loss of resources).

Both self-regulation models combine metacognitive strategies, cognitive strategies and affect. Specifically, they both include the following components: task analysis, self-evaluation, self-reactions and self-motivational beliefs. In both models, the learner or test-taker analyzes the task and assesses his / her own ability to deal with the task demands that may result in positive or negative self-reactions or emotions. These emotions in turn affect task performance. In the control model the underlying assumption is that individuals "self-regulate their behavior in terms of two basic priorities" (Boekaerts & Niemivirta, 2000:428). This assumption of two parallel processing modes is the main difference between the two models. In the control model, individuals either want to increase their knowledge and skills or they want to maintain or prevent damage to their well-being. One criticism of the control

Figure (4): The model of adaptable learning
(Boekaerts & Niemivirta, 2000:429)



model is that it does not include the possibility of the impact of coping strategies on knowledge and skills. It does not show that negative appraisals may affect metacognitive control. Undoubtedly, self-regulation plays a central role in influencing performance on tests and key elements from these two self-regulation models will be used in positing a model for emotional regulation in test situations. The following section examines self-regulation while taking tests focusing on the role of affect.

2.5.3 Emotional Regulation in Test Situations

Emotions and emotional regulation in education and specifically in test situations have been largely neglected except for test anxiety that has been researched extensively (Pekrun et al., 2002; Schutz & DeCuir, 2002; Schutz & Davis, 2000). Emotional regulation includes processes that focus on “monitoring, evaluating and modifying our emotional experiences” and “involves flexible, situationally responsive and performance-enhancing strategies” (Schutz & Davis, 2000:243). Thus, emotional regulation is defined in terms of processes and strategies. There have been recent approaches in investigating emotions in education within a process approach. Scherer (2000:75) proposed a model that included the dynamic and fluctuating nature of emotion processes and “discrete language labels referring to steady states” i.e. emotions as static variables. Schutz and Decuir (2002:128) define emotional regulation as “interdependent and interrelated nonlinear processes” and “a particular emotional experience involves dynamic and continually changing processes in which the individuals attempt to make meaning out of their particular transaction.” Schutz & Davis (2000) and Schutz et al (2002) proposed a model of emotion and emotional regulation during test-taking that included four conceptual domains: cognitive-appraising processes, task-focusing processes, emotion-focusing processes and emotions experiences. The model includes strategic behavior which consists of test and study strategies as a separate component.

Schutz & Davis (2000) differentiate between strategies and processes. Strategies are specific actions, behaviors or techniques that test-takers apply consciously or unconsciously while taking a test. They define process as “a more

general term related to changes that occur in the human system over time... A process may or may not involve strategic activity (p.246).” They add that these processes are used as strategies for the regulation of emotions. Therefore, emotional regulation during test-taking consists of the emotions that are experienced during a test and the processes that are applied to deal with them. These emotional regulation processes are now described in detail.

Emotional regulation starts with a comparison between goals we attempt to attain and the situation at that particular in time. This appraisal process or cognitive-appraising processes (Purpura (1999) and Bachman & Palmer (1996) as previously mentioned refer to this appraisal process as assessment or metacognitive strategy use) lead to the emergence of emotions or emotional experiences. Different types of appraisals lead to different emotions. For example, for test anxiety to emerge, we must first judge the test as being an important and a relevant goal. If we have little or no confidence in completing the test tasks, anxiety emerges. The four key cognitive appraisal processes that influence the types of emotions that occur are:

- the importance or goal relevance of the test: if the test is not judged as important, then emotions are not likely to emerge;
- goal congruence of the test: if what is happening during the test is judged as helping reach one’s goal, then pleasant emotions such as happiness or optimism are likely to emerge and if not, then unpleasant emotions such as anxiety or fear are more likely to occur;
- control: judgment of the test-taker of who is in control or judgment about the cause of what is happening during the test. If the test is judged to be relevant but it is not going well and someone is to blame, then anger is likely to emerge. Pride is likely to emerge if the test is judged to be important and is going well;
- dealing with problems that occur during the test (ability to cope): judgment of the test-taker about his / her ability to handle what is happening during the test. If the test is judged to be relevant but is not going well and test-taker is not able to handle the test, then anxiety is likely to emerge. If the test is judged to be relevant but is going well and test-taker is able to handle the test, then challenge or hope is likely to occur.

Appraisals test-takers make reflect the way they view the outside world. This in turn implies that emotional regulation can be controlled. Thus, if the appraisal is

changed, then the emotional experience can also be changed. Zeidner (1998) discussed regulation of test anxiety that succeeded over time using interventions that included identifying anxiety-facilitating beliefs in the treatment.

Task-focusing emotional regulation processes were categorized by Schutz & Davis (2000) as:

- task-focusing activities such as identifying the key items in a question, managing time, reading directions and checking answers “to keep focused on the test and away from potentially disruptive negative thoughts ...” (p.248);
- tension reduction such as self talk or slow breathing;
- importance reappraisal processes such as “attempts to keep the importance of the test in context or to emphasize the positive aspects of the test, ...” (p.248).

The third domain of the model included emotion-focused processes where test-takers’ attention is focused on their emotions and away from the test. Examples of emotion-focused processes are:

- wishful thinking: “involves thoughts like hoping the problem will just go away or hoping the teacher will not count the test”;
- self-blame: “involves criticizing ourselves about our handling of the test or our preparation for the test”;

Here the focus is on feelings about the test and not on the actual test task. However, there is some overlap here with emotions since wishful thinking is an emotion and cannot be seen as a process.

Another term that has been used in emotion regulation during test-taking is coping. In 1995, Zeidner reviewed the literature on students’ coping with test situations. Coping was defined as a person’s cognitive and behavioral efforts in dealing with a situation that is appraised as stressful or threatening such as an important examination. Thus, it is the ability to control the effects of negative emotions and perform well even when there is emotional distress. Individuals with low skills in coping are likely to quit at first signs of emotional distress. The coping process consisted of a (1) primary appraisal: evaluating a situation and determining whether it is threatening or challenging; a (2) secondary appraisal: which included context-specific judgments such as difficulty of the test, expectations of success and control over outcomes; and (3) specific coping responses or strategies. The research reviewed focused on exploring coping behaviors and emotions across four

stages of a stressful examination encounter: the anticipatory stage, the confrontation stage, the waiting stage and the outcome stage. The anticipatory stage is the phase before the examination while the confrontation stage is the phase where the test-takers actually take the examination. The waiting stage refers to the phase after having taken the examination until the results are announced while the outcome stage is the phase after the grades are announced. Zeidner found “very few studies have objectively assessed how students actually feel and think about the exam at this critical stage” (p.125) i.e. the confrontation stage. He mentioned that test-takers use a variety of task-oriented and palliative coping strategies during a test but he did not specify what they were. In general, the research he reviewed showed that test-takers use a variety of task-oriented and emotion-oriented forms of coping in the anticipatory, waiting and outcome stages of exam situations. Zeidner also found that coping strategies before exams had a modest impact with respect to affective outcomes but had little effect on exam performance. There was no consensus in the research about “which coping strategies are most effective and adaptive in promoting positive outcomes in exam situations” (p.132). Individual differences were found in reacting to an evaluative situation. Thus, coping strategies varied from one person to another and also “may change over time in order to manage both short- and long-term effects of a stressful examination (p.132).” He stated that more research was needed in order to find out how a coping strategy relieves emotional distress and solves problems. He recommended that further research on coping strategies in examination situations should include situational and personal variables because specific coping strategies used by a test-taker while taking a test have not clearly delineated. Thus, Zeidner concluded that further research was needed on coping or emotional regulation during test-taking. Zeidner defined coping as a process that included appraisal, emotions and specific coping strategies or behaviors. He did not propose a model that described coping in exam situations.

Schutz & Davis (2000) noted that most of the research on test-taking has focused on emotional regulation from a coping perspective and “more pleasant emotions have not been investigated from the coping point of view” (p.253). They ask questions about what are the processes that regulate positive feelings and how would test-takers keep from getting overconfident or too pleased with themselves.

Therefore, from the review of the literature it was found that emotional regulation consists of processes and positive and negative emotions with test-taking

strategies as a separate component. Coping refers to strategies or processes used to regulate negative emotions and more research is needed on strategies or processes used to regulate positive emotions.

2.5.4 Research Questions 2 & 3

In light of the self-regulation literature reviewed above, Research Question 2 is modified to the following:

What are the test-taking strategies and emotional regulation processes used and emotions experienced by Egyptian adult EFL learners in a specific context when taking an English language placement test?

Research Question 3 is posed as follows:

What is the relationship between test-taking strategies, emotional regulation and test performance for Egyptian adult EFL learners in a specific context when taking an English language placement test? Further questions that focus on the relationship between test-taking strategies, emotional regulation and test performance are:

- 3.1 Do test-taking strategies vary across different levels of test performance?
- 3.2 If so, how do test-taking strategies differ across different levels of test performance?
- 3.3 Does emotional regulation vary across different level of test performance?
- 3.4 If so, how does emotional regulation processes differ across different levels of test performance?
- 3.5 Is there a relationship between test-taking strategies selected and emotional regulation?
- 3.6 If so, does the relationship between test-taking strategies selected and emotional regulation differ across different levels of test performance?

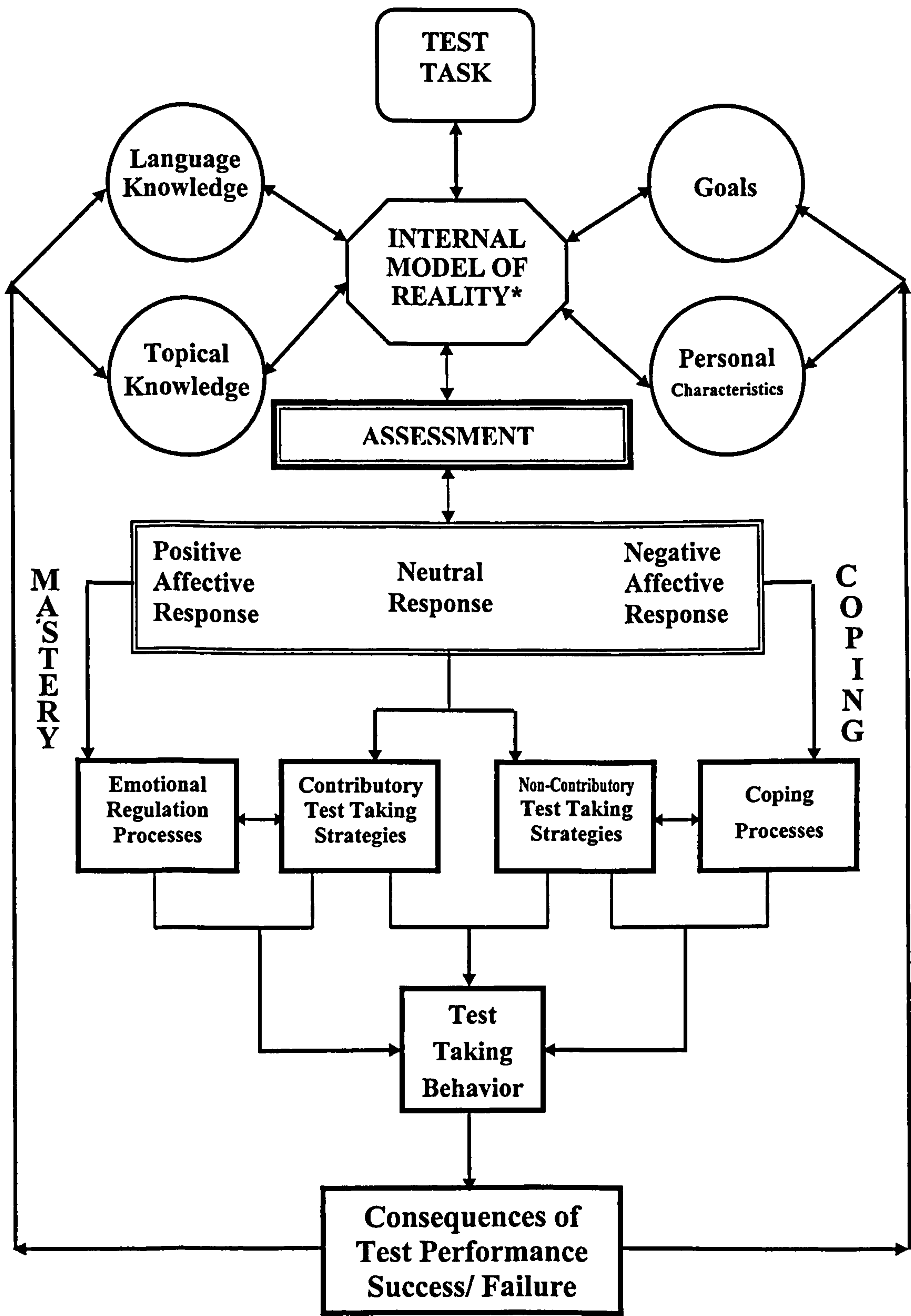
The following section discusses a language testing processing model that describes the test-taking and emotional regulation processes that a test-taker applies while taking a test.

2.6 PROPOSED LANGUAGE TESTING PROCESSING (LTP) MODEL

In order to address the research questions within a language testing processing framework, the metacognitive strategies component of Bachman & Palmer's model is modified based on the results of Purpura's (1999) study and incorporating elements of the Model of Adaptable Learning (Boekaerts & Niemivirta, 2000) and Schutz & Davis's (2000) emotional regulation model. From Purpura's study the three metacognitive strategy elements are replaced by an 'assessment' component and test-taking cognitive strategies that contribute positively to test performance and test-taking cognitive strategies that contribute negatively to test performance are included in the model. From Boekaerts & Niemivirta's model, the two parallel processing routes: mastery and coping are included in the LTP model. In the LTP model, mastery means that the test-taker is in control of the test-taking situation, while coping occurs when the test-taker experiences a negative affective response. Emotional regulation of both pleasant and negative emotions is included as a component in the proposed model. From the emotional regulation model, emotional regulation processes and behaviors are incorporated. A test-taker's internal model of reality is based on four sources of information: language knowledge, topical knowledge, goals and personal characteristics. These four sources contribute to the appraisal process which consists of analysis of the test task and the assessment component (metacognitive strategy use). The new model is shown in Figure (5). In assessing ability to perform a test task (the whole test, a section or a test item), the test-taker judges personal capabilities such as knowledge, skills, strategies or personality traits balanced against perceived personal weaknesses in the particular testing context. The test-taker makes comparisons between his / her goals and the test situation. If the test is relevant to the goals of a test-taker, then positive or negative emotions emerge. These assessments involve both feelings and cognitions or cognitive strategies. The results of these assessments may or may not lead to a discrepancy between the requirements of the test task and language knowledge, topical knowledge, goals and personal characteristics. If there is no discrepancy and the test is judged as not important to the test-taker's goals, no intense positive or negative emotions are experienced and test-takers apply contributory and non-contributory cognitive strategies to the test-taking process. If

there is a discrepancy and the test is judged as being an important and relevant goal, it may lead to positive emotions where test-takers apply contributory cognitive test-taking strategies which lead to successful test task performance. Contributory strategies are those that contributed to the choice of the correct answer (Nevo, 1989). Emotional regulation processes here include task focusing processes such as monitoring time, self-encouragement and keeping negative or disruptive thoughts away. A discrepancy may also lead to negative emotions or distress where test-takers attempt to cope by applying non-contributory cognitive test-taking strategies which may lead to unsuccessful test task performance. Emotional regulation or coping processes include tension-reduction strategies such as slowing down, self-talk and taking deep breaths. Test-takers may disengage from the actual test task and focus on their emotions and thoughts about their performance. These emotion focusing processes or coping processes may include “daydreaming about how things could be different, self-blame or self-criticism about how one is doing” (Schutz et al: 2002:320). The test-taker’s behavior during the test-taking process is the outcome of the selection and adaptation of cognitive strategies and of processes for managing emotions. Test-takers may increase or decrease effort, persist in completing the task, give up or seek help (ask for answers or even attempt to cheat during the test). The consequences of test performance are either successful or unsuccessful experiences which are a powerful source of knowledge about one's own capabilities. The perception that one's performance has been a failure may contribute to the expectation that future performances will also be unsuccessful and success contributes to the expectation of proficient performance in the future. External experiences or assistance also contribute to a test-taker's internal model of reality. The LTP model is a theoretical model and needs to be validated by data. It remains to be determined whether the model would be useful in explaining or describing data obtained from test-takers while they took a test.

Figure (5): The proposed language testing processing model
(after Bachman & Palmer, 1996)



* Frame of reference

2.7 DEFINITION OF TERMS

Test-taking strategies are the cognitive strategies used to complete the test task and respond to test items. Emotional regulation includes an assessment component, emotional regulation processes and emotions. The assessment component involves the application of metacognitive strategies which are strategies used by test-takers to judge their capabilities balanced against perceived personal weaknesses for each test task. Emotional regulation processes include task-focusing and emotion-focusing or coping processes applied by test-takers to handle positive or negative emotions during the test-taking process.

2.8 SUMMARY

In conclusion, an extensive literature review of language testing models, language test-taking strategies, language self-assessment, research in affect and test performance including language anxiety, self-regulation and emotional regulation during test-taking resulted in the identification of gaps in the literature that need to be addressed. These gaps are:

- in order to apply a modified Bachman & Palmer (1996) model i.e. the LTP model, the first step required is to define the construct of language proficiency. Because there is no generic definition of language proficiency, the construct of language proficiency needs to be empirically defined for this particular Egyptian context.
- the different test-taking strategies used and test-takers' affective responses by both high ability and low ability adult EFL test-takers while completing different test tasks: writing, reading & listening using both think alouds and checklists / questionnaires.
- a language test processing model that describes the test-taking process and includes metacognitive and cognitive strategies and emotional regulation processes.

Identifying these gaps, in turn, led to the three research questions which are the focus of this study and which are investigated within a specific Egyptian EFL context. The purpose of the first research question is to establish a contextualized

definition of language proficiency in the specific Egyptian EFL context which is the first step required in any language test processing model. The second and third research questions investigate the test-taking strategies and emotional regulation processes used and emotions experienced by Egyptian EFL adult learners during test-taking. Differences in test-taking strategy use by low and high ability learners are also examined. The literature review also led to the postulation of a model that attempts to describe the interplay of metacognitive and cognitive strategy use and affective factors during the test-taking process. This model is further discussed and developed in Chapter 5 dealing with Research Question 2. The following chapter describes the research design and methodology followed in this study.

CHAPTER 3:

RESEARCH DESIGN

3.1 OVERVIEW OF RESEARCH DESIGN

The goal of this study is to establish a contextualized definition of language proficiency in a specific Egyptian EFL context, to identify learners' test-taking strategies and emotional regulation processes used during test-taking and to explore the relationship between test-taking strategies, affective factors and performance on a language test. This goal falls within the area of assessment validation. The research approach and methods have been chosen to fit the focus of the study which is concerned with the effects of cognitive and affective factors on language test performance of Egyptian adult learners in a specific context. Historically, validation research in language assessment has been dominated by the positivist / psychometric approach and only in the last few years has the language assessment community begun to extend the discussion of validity beyond the traditional paradigm and formulate new perspectives. Moss (1996) in particular examined validity within the interpretive paradigm. She contrasted the traditional approach with the interpretive "to highlight the taken-for-granted practices and perspectives of each approach and how, taken together, they can provide a more textured and productive view of the social phenomena we seek to understand" (p.22). In line with this shift in the field, the overall research approach of this study is eclectic combining elements of both the positivist and interpretive paradigms.

The interpretive paradigm is "characterized by its concern for the individual" and "efforts are made to get inside the person and to understand from within" (Cohen & Manion, 1994: 36). In order to define the construct of proficiency in the specific Egyptian EFL context, it was investigated from both the learners' and the teachers' perspectives in the specific context. Interpretive researchers "begin with individuals and set out to understand their interpretations of the world around them" (Cohen & Manion, 1994: 37). No universal or international definition of language proficiency is adopted, for example, viewing language proficiency in terms of the four skills of listening, speaking, reading and writing; rather it is generated from within the context. Investigating the cognitive and affective factors of learners and relating

them to language test performance is firmly located within the interpretive paradigm. This research uses a qualitative methodology to seek understanding of language test performance, however, the approach is also complemented by the use of quantitative data collection procedures, data collection tools and data analysis. Sampling was done using the positivist approach and a Likert-type questionnaire was also one of the tools used. The qualitative data obtained were quantified and frequency counts were reported. The Likert-type questionnaire data are statistically analyzed.

The two approaches of interpretivism and positivism have been combined resulting in a mixed model research design or a pragmatic approach. The pragmatic approach which characterizes this study involves mixing quantitative and qualitative approaches in different phases of the research process.

3.2 RESEARCH PARADIGMS

Paradigms are defined as a basic set of belief systems that guide researchers (Guba & Lincoln, 1994). They are the basic assumptions and practices shared by a particular community engaged in scientific endeavor. A paradigm is a framework that determines how the members of a community define problems for inquiry and defines the theories, methods and techniques used to solve these problems. Thus, a paradigm determines the way the members of a community view the world and discover the 'truth' of the world (Usher, 1996). In language assessment, the community consists of test designers, researchers and those who administer and interpret the meaning of tests. Kuhn (1970) claimed that paradigms are developed and then are replaced when they cannot explain anomalies or problems. Thus, in periods of transition and change competing paradigms may exist simultaneously i.e. both research and practices in assessment are conducted or operate within two paradigms (in recent years alternative language assessment has emerged in parallel to the traditional approach to language assessment (Lynch, 2001)). Major readjustments in thinking and the traditional patterns continue to operate concurrently. Over the last two decades, the positivist approach to educational research has declined, whereas the interpretivist approach has been on the increase. As a result of this "interpretive turn," the old debate between positivism and interpretivism (and central to the quantitative-qualitative debate) has diminished in intensity ..." (Howe, 1998: 13).

Looking at the two paradigms as located at the opposite ends of a continuum, the following differences on certain dimensions between each paradigm are clearly drawn. In the positivist paradigm an objective external reality is assumed to exist that can be observed and measured and research can converge on a true state of affairs. It assumes that this reality is governed by laws and mechanisms and can be investigated independent of who, when and how it is being examined. When influence in either direction (threats to validity) is recognized steps are taken to reduce or eliminate it (Schwandt, 1994). Within positivism, researchers are likely to focus on collecting large amounts of data in order to be able to generalize to a large number of situations. Introspective data and studies involving only small groups cannot be relied upon. Interpretivists maintain that "reality is perceived as a construct of the human mind such that there can be different interpretations of what is real" (Allan, 1996: 103) and consequently there are multiple realities which exist to fit contexts and purposes of human acts. This reality is dependent on the inquiry and attempts to know are inherently subjective. Research is influenced by the values or framework an investigator uses. Research is regarded as a process and findings are literally created as the investigation proceeds (Schwandt, 1994; Tashakkori & Teddlie, 1998).

Guba & Lincoln (1994) have emphasized the differences between the paradigms stating that one precludes the other and that trying to combine the two approaches is doomed to failure due to different underlying philosophies. Smith (1983:12) described the incompatibility of the two paradigms as follows:

"one approach takes a subject-object position on the relationship to subject matter; the other takes a subject-subject position. One separates facts and values, while the other sees them as inextricably mixed. One searches for laws, and the other seeks understanding. These positions do not seem incompatible."

To counter this argument, Howe (1998) posited a different paradigm: pragmatism where both qualitative and quantitative methods are compatible and researchers could make use of both of them in their research. Guba & Lincoln (1994) reiterated that research methods are secondary to questions of paradigm. Both quantitative and qualitative methods can be used with any research paradigm,

however, it is the paradigm which defines our basic belief system and the way we view the world. This means that interviews would be used in positivist research and questionnaires in purely interpretivist studies. Pragmatists consider the research method to be secondary to the research question itself and the worldview underlying the method is not seen as important. Howe (1998:13) states “but why should paradigms determine the kind of work one may do with inquiry any more than the amount of illumination should determine where one may conduct a search? ... Eschewing this kind of tyranny of method ... is the hallmark of pragmatic philosophy.” The following Table (4) taken from (Tashakkori & Teddlie, 1998:23) compares the three paradigms being discussed: positivism, pragmatism and interpretivism in terms of methods, epistemology, axiology, ontology and causal linkages. The table shows pragmatism as rejecting the either-or point of view or the incompatibility position and includes both.

Table (4): Comparison of positivism, pragmatism & interpretivism

Paradigm	Positivism	Pragmatism	Interpretivism
Methods	Quantitative	Both quantitative and qualitative	Qualitative
Logic	Deductive	Both deductive & inductive	Inductive
Epistemology	Objective point of view. Knower and known are separable.	Both objective & subjective points of view.	Subjective point of view. Knower & known are inseparable.
Axiology	Inquiry is value-free.	Values play a large role in interpreting results.	Inquiry is value-bound.
Ontology	Objective external reality.	Accept external reality. Choose explanations that best produce desired outcomes.	Multiple social realities that are products of human intellects and may change.
Causal linkages	Real causes temporally precedent or simultaneous with effects.	There may be causal relationships, but we will never be able to pin them down.	All entities simultaneously shaping each other. It's impossible to distinguish causes from effects.

Although interpretivism is more usually associated with qualitative methods, interpretivist research can also include quantitative methods i.e. the use of quantitative methodology to analyze what is essentially qualitative data. The paradigm wars, or debates, are now over (During the 1980s and early 1990s researchers would defend their own methodological position and attack the other side. This debate regarding the superiority of one or the other of the two major paradigms became increasingly unproductive and was finally abandoned.) and pragmatic or eclectic researchers use mixed methods which have elements of the two approaches: quantitative and qualitative. Many researchers use the method that is most appropriate for their study, rather than rely on one method exclusively (Tashakkori & Teddlie, 1998). The decision to use a particular method depends on the research question.

The research cycle can move from facts and observations through induction to generalizations or theory or from theory through deduction to hypotheses or prediction of events or outcomes. Research can start at any point in the cycle and in mixed model studies both types of methods are used.

Pragmatists challenge the incompatibility thesis that researchers can be both objective and subjective in investigating a research question. Using both quantitative and qualitative methods inevitably leads to embracing both objective and subjective points of view.

Tashakkori & Teddlie (1998:26) describe the role of value systems to pragmatists as follows:

“...pragmatists decide what they want to research, guided by their personal value systems; that is, they study what they think is important to study. They then study the topic in a way that is congruent with their value system ... and appropriate for finding an answer to their research question. They also conduct their studies in anticipation of results that are congruent with their value system.”

This approach describes the way many researchers in the social sciences conduct their studies.

Guba & Lincoln consider ontology, or nature of reality and causality, to be the main differences between positivism and interpretivism. Pragmatists agree that there is an external reality however, they “deny that “Truth” can be determined once

and for all” (Tashakkori & Teddlie, 1998:28). They also believe that knowledge cannot be separated from beliefs and interests. Thus, they are not sure that one explanation of reality is better than another.

Considering the previous arguments, pragmatism is appealing because it presents a paradigm that includes the use of mixed methods and presents a practical research philosophy: “Study what interests and is of value to you, study it in the different ways that you deem appropriate, and use the results in ways that can bring about positive consequences within your value system.” (p.30)

There is considerable overlap between pragmatism and the postmodernist context or approach of educational research. The basic core of postmodernism is doubt in what is known and the ways of knowing. Within postmodernism, all approaches to research reflect researchers’ cultural beliefs about the world and it questions absolutes and the argument that following a correct method of systematic observation leads to the “Truth”. Postmodernism is not anti-science but challenges the view that there is a “determinate world that can be definitively known and explained” (Usher, 1996:25). Yet, a postmodernist approach does not simply adopt the interpretive research tradition since it does still embrace the positivist framework. It rejects the dichotomy between positivism and interpretivism and suggests an alternative which challenges the dominant discourse in its different forms. Postmodernists question whether knowledge is generated through empirical observation and experiment or whether it (i.e. knowledge) is constructed in different ways.

Interestingly, since this study focuses on language assessment, language plays a central role in postmodernism. Postmodernism views knowledge generation as occurring through language and discourse. Language is not seen as a reflection of social reality or as a tool for conveying meaning of an external reality. Language produces reality, it is seen as both the carrier and creator of social reality. Thus, no form of knowledge can be separate from the language and discourse operating within a culture. “The structures, conceptuality and conventions of language, embodied in discourses and texts ... govern what can be known and what can be communicated” (Usher, 1996:27). Knowledge then is relative and is always partial. Thought and experience are dependent on socio-cultural contexts and practices. Language is how social organization is defined and how subjectivity is constructed. Language is seen as varying discourses: different ways of giving meaning and viewing the world.

Language constructs an individual's subjectivity in a specific context and it is not the result of one's individuality. Since individuals are subject to multiple discourses, their subjectivity is changeable and not rigid or fixed (Richardson, 2000 in Denzin & Lincoln). The implication of this view to researchers and writers is that we must "understand ourselves reflexively as persons writing from particular positions at specific times" (p.929).

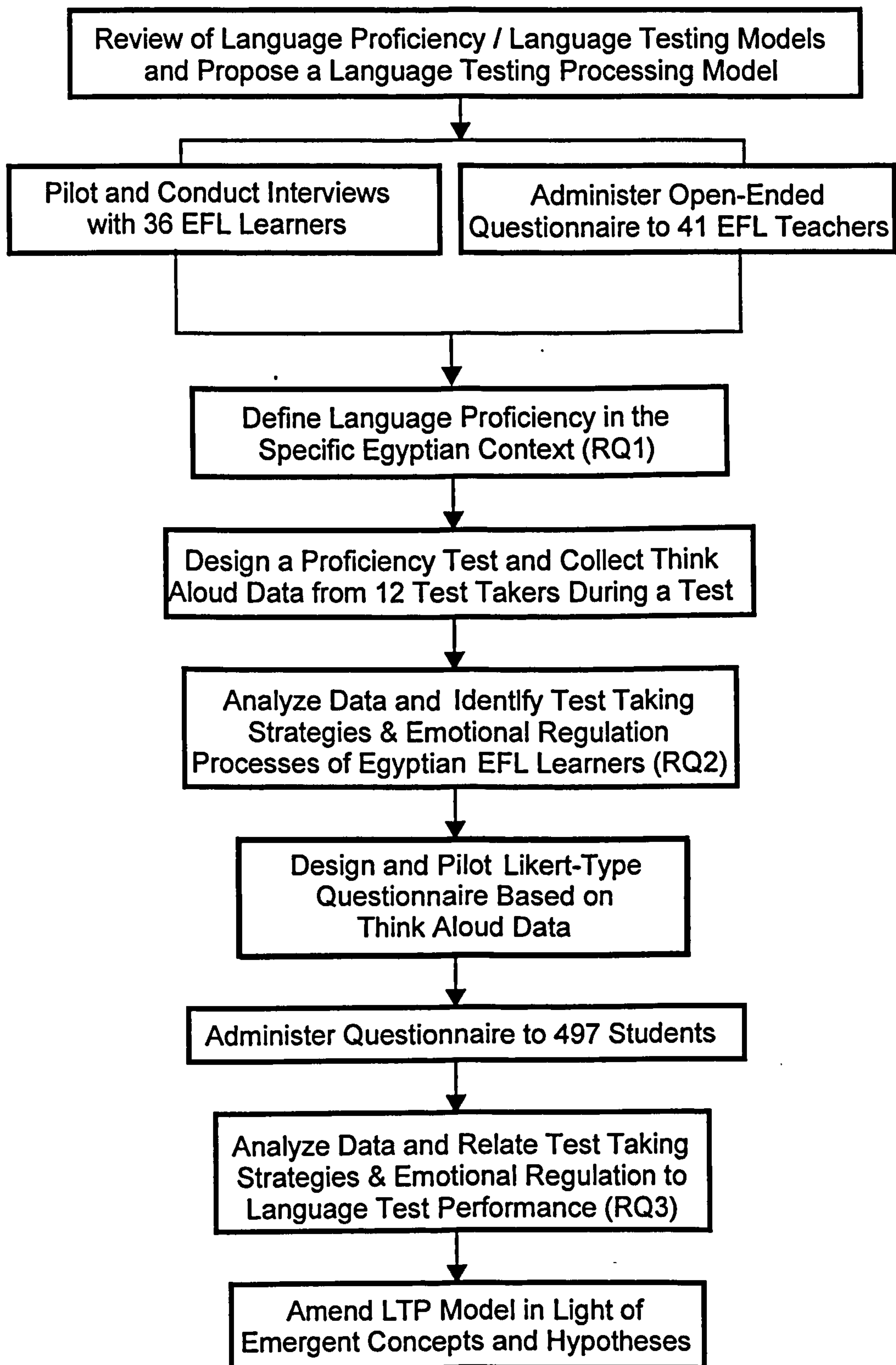
In order to determine the paradigm in which this study is located, it is important to examine the underlying assumptions of the research itself in relation to each paradigm. The following Figure (6) illustrates the design of the study and shows the different stages involved. The diagram includes the sample size of respondents who participated in the various data collection phases. Figure (6) also shows the sequence followed in investigating the three Research Questions: RQ1, RQ2 and RQ3. RQ1 focused on establishing a contextualized definition of language proficiency at CACE, AUC and the purpose of RQ2 was to identify learners' test-taking strategies and emotional regulation processes used during test-taking. RQ3 focused on exploring the relationship between test-taking strategies, affective factors and performance on a language test.

3.3 RESEARCH PARADIGM OF THE STUDY

Research Question 1 seeks to define the construct of foreign language proficiency in a specific Egyptian EFL context. Within this context, Research Question 2 seeks to identify the test-taking strategies and emotional regulation processes used and emotions experienced by adult EFL learners when taking an English language placement test and Research Question 3 examines the relationship between test-taking strategies, emotional regulation and test performance for adult EFL learners taking an English language placement test.

In Research Question 1, the construct of language proficiency is seen to exist as an external reality, however it is generated and defined based on individuals' perceptions and beliefs about language use in this particular context. The focus of validity is not on the test but on the meaning of the scores. This study investigates construct validity from a pragmatic perspective where meanings are not predetermined but interpretations are made from a particular assessment context. The definition of language proficiency is generated from the data obtained through semi-structured interviews with learners and teachers in an Egyptian EFL context.

Figure (6): Research design



There were no preconceived assumptions of what language proficiency is for these individuals.

For Research Questions 1, 2 & 3 at the ontological and epistemological levels, the study is conducted within the pragmatist paradigm using both quantitative and qualitative methods. The research questions or issues in this study are investigated from a pragmatist point of view using a combination of qualitative and quantitative methods. The researcher does not regard “the researched as data objects and sources of data” rather the individuals included in this study are viewed as “valid interpreters of their own social world and social actions” (Byrne, 2002: 146).

To further elaborate, in this study a distinction is made between subjective experiences such as perceptions, emotions and thoughts and the objective realities of thoughts e.g. “content of books, theoretical systems, ...” (Schutz & Decuir, 2002:126). When investigating the definition of language proficiency and affective factors in test-taking, not only do these constructs need to be studied but also the context “from which those concepts emerge and the process and meaning of those concepts to the individuals involved (p.126)” has to be taken into consideration. Context is seen to be a significant factor when conducting research and interpreting data. Subjective experiences and objective realities of our thoughts are subject to continual change and a pattern identified now may not be the same in the future. Human action is goal directed and these goals emerge within a context. Individuals interact within this context and these goals are manifestations of the context from which they emerged. Human beings have the capability of reflecting on their own context and the objective realities of their thoughts. These perceptions are developed through reflection and constitute knowledge about subjective experiences and objective realities of thoughts. This “knowledge represents the basis from which appraisals are made about the world (p.126)”. Thus, an understanding of the context must be developed in order to define the construct of language proficiency and for us to be able to understand affective factors in taking tests. The nature of knowledge is changeable and what we know is “subject to change based on changes in the external world or in the way we construct our understanding of the external world (p.126)”.

At the methodological level, positivists use quantitative methods and research is conducted using the experimental approach where variables are controlled and manipulated. The interpretive paradigm uses qualitative methods and a

nonexperimental set of procedures that have been referred to as hermeneutical or dialogic. Hermeneutics is

“a research process in which the researcher forms interpretations, or constructions, from a close understanding of the data (observation notes, interview recordings, etc.). This initiates a cycle (potentially neverending: hermeneutic circle”) of interpreting these constructions and then refining and forming new constructions. Dialogic refers, basically, to methods that involve an interaction with participants in the research setting, a dialogue that encourages the participants to develop an understanding of what is being researched, or learned, in their own words, their own terms.”

(Hamp-Lyons & Lynch, 1998: 261)

In a pure interpretivist paradigm qualitative methods are only used in data collection and in data analysis. However, this is where this study deviates from a pure interpretivist approach and incorporates the use of quantitative methods in data collection and data analysis. This research uses a qualitative methodology that is “characterized as beginning with the individual rather than the group” (Banerjee & Luoma, 1997: 276) and the data obtained are verbal rather than numerical data. In conducting the semi-structured interviews used to collect data for Research Question 1, which focused on defining the construct of language proficiency, a quantitative (positivist) approach to sampling was used and the learners were randomly selected taking into consideration three background variables: gender, age and proficiency in English. The open-ended questionnaire used in the same phase of the research was administered to the total population of teachers. The questions and procedure of the semi-structured interview were organized in advance with some flexibility to probe in specific areas which was also predetermined. The qualitative data obtained for two research questions of the study (Research Questions 1 & 2) were processed and analyzed using frequency counts. In Research Question 3, the Likert-type questionnaire was designed based on data generated from the verbal protocols. The data obtained from the Likert-type questionnaire were analyzed quantitatively.

The gradual development of research designs that incorporate quantitative and qualitative methods was, to a large extent, due to the popularization of the use of triangulation methods. Triangulation involves the use of combined data sources to study a particular research problem. One of the types of triangulation is methodological triangulation which is classified into two types: within methods

triangulation (using multiple qualitative or multiple quantitative methods) and across methods triangulation (using both quantitative and qualitative methods) (Tashakkori & Teddlie, 1998). The basic assumption of triangulation is that any bias inherent in one data source would be neutralized when combined with other data sources (Creswell, 1994). Besides triangulation being an important reason for using both qualitative and quantitative methods, mixed methods provide richer detail and add scope and depth to a study. Other reasons for using mixed methods are that they: are complimentary and “overlapping and different facets of a phenomenon may emerge” (Creswell, 1994: 175); initiate new lines of thinking leading to contradictions and fresh insight (Allan, 1996); and are developmental where the methods are used sequentially to provide a deeper analysis. In a sequential mixed method design, the researcher conducts separate quantitative and qualitative phases.

In this study, qualitative methods are used to investigate the construct of language proficiency in Research Question 1 and to further explore validity through the use of verbal protocols which is a qualitative method used in language test validation where the test-takers verbalize their thoughts while completing the test in Research Question 2. In order to corroborate the findings from the qualitative phase, a subsequent quantitative phase of the study was designed and a Likert-type questionnaire was used for Research Question 3. This is a common type of sequencing, because in most quantitative survey research, quantitative instruments are developed based on analyzing qualitative data. This process of sequencing qualitative / quantitative data collection can be used iteratively going through several cycles (Tashakkori & Teddlie, 1998; Allan, 1996). Therefore, the research design is eclectic or pragmatic where aspects of the two paradigms are drawn upon as appropriate at several stages of the research process. As stated by Creswell (1994: 178) “this approach adds complexity to a design and uses the advantages of both the qualitative and quantitative paradigms.”

3.3.1 Case Study

According to the following definition, this research can also be considered as a case study

“in which the researcher explores a single entity or phenomenon (‘the case’) bounded by time and activity (a program, event, process, institution, or social group) and collects detailed information by using a variety of data collection procedures during a sustained period of time”

(Creswell, 1994: 12)

One goal of a case study is to study a phenomenon in its natural context and develop an understanding of a phenomenon as the participants view it (Yin, 1994) or as they experience it. This research investigates the construct of language proficiency in a particular context (the English language program of the Center for Adult and Continuing Education (CACE) at the American University in Cairo). This study is based on Egyptian adult learners’ perceptions in this particular context. This thesis also investigates test-taking strategies and emotions regulation experiences during a test as reported by the respondents themselves. Since the definition of a case study does not specify a particular paradigm or method, this study which uses a variety of methods at different stages, fits the definition and can be classified as a case study.

3.4 DATA COLLECTION INSTRUMENTS

The overall design of the study is depicted in Figure 6. The data collection instruments used in the study are (see Appendices 2, 3, & 4 for actual instruments):

- semi-structured interview;
- open-ended questionnaire;
- think aloud verbal protocols;
- Likert-type questionnaire.

The following sections will describe the rationale, design and piloting each of the data collection instruments.

3.4.1 Semi-Structured Interview

3.4.1.1 Rationale

Although interviews are classified into quantitative, closed-ended and qualitative, open-ended, there is a continuum ranging from structured to unstructured interviews. The semi-structured interview which is located in the middle of the continuum was selected as the most appropriate data collection tool since it allows

depth to be achieved in investigating learners' perceptions regarding language proficiency. In a semi-structured interview there is a one-to-one interaction between the researcher and the individual and it provides the opportunity for ideas to be followed up, to probe and expand on interviewee's responses and to ask for further clarification. A technique that has grown in popularity within the methodology of interviewing is the focus group interview. In this study the semi-structured interview is preferred to the focus group interview because of two main reasons. Based on the researcher's experience in this context three key background variables must be considered when sampling subjects: proficiency level, gender and age. One of the difficulties faced would be to organize focus groups (usually consisting of eight to ten interviewees) ensuring that these background variables are all represented. All the learners are adults who come to attend classes in the evening, twice a week. Each class has about 22 students, therefore, to arrange for focus groups, intact classes at each proficiency stage cannot be used and scheduling focus groups and getting the right mix of individuals would be very difficult. Another risk faced is that social pressure will influence responses (Cohen & Manion, 1994; Krathwohl, 1998), in particular, those of females in this context in Egypt where females are more likely affected by peer pressure (this is based on the researcher's observations and experiences in the Egyptian context).

3.4.1.2 Design of the Instrument

The questions in the Student Interview Protocol (SIP) were set prior to the interview and were designed based on the following aspects of language proficiency: to find out learners' motivation for studying English, contexts of language use and their expectations in terms of desired terminal language objectives. The actual questions are illustrated in Appendix 2 (Student Interview Protocol).

3.4.1.3 Piloting of the Student Interview Protocol

The SIP was designed based on the following objectives: to find out students' motivation for studying English, the different contexts where they use English, their self-perception of their current English language ability and the desired level of proficiency they want to attain. The questions initially formulated to match these objectives were as follows:

- Why are you studying English?
- What are the situations where you use English?
- What level of proficiency are you aiming for?
- How long do you intend to study English?
- How would you evaluate your listening, speaking, reading and writing abilities?

The interview questions were piloted on a male subject at the basic level in the 23+ age group (a copy of the transcript of the interview is included in Appendix 5). The student explained his reason for studying English and stated that he had studied English before. He described his use of English in general and thus, the researcher had to elicit more specific responses regarding the functional use of English. A better question would have been to ask about the difficulties faced in language situations. In a study by Bachman & Palmer (1989) they concluded that the most effective question type to elicit students' self-perceptions of their current language ability was one which asked about their perceived difficulty in different areas of the language. The student was also unable to describe the level of proficiency he was aiming for and required more focus when asked to evaluate himself in terms of the four language skills. Since the aim was to establish a definition of proficiency, it was decided to move away from getting students to describe their own language proficiency to focusing on defining the language of a person who is at the highest stage of proficiency. Based on these results, the SIP was modified as follows:

- Why do you want to study English?
- Do you use English at work / at home? How?
- Have you ever been in an English language situation you found difficult? What were the difficulties?
- Describe the language of someone who is fluent or who is at the highest stage of proficiency.

These interview questions were piloted a second time on a male subject at the advanced level in the 23 - 35 age group (a copy of the transcript of the second interview is included in Appendix 5). Two questions required rephrasing. Question

two was changed to “Where do you use / are you going to use English?” in order to focus on the specific contexts of use of the language. Question three was rephrased to avoid limiting students' responses to a specific situation and to emphasize the difficulties they face in general in using the language. The student found it difficult to think of a specific situation. The final SIP was as follows:

- Why do you want to study English?
- Where do you use / are you going to use English?
- Describe the difficulties you face in using English.
- Describe the language of someone who is at the top level of proficiency.

Because, the SIP was based on specific objectives and the questions were used to guide and focus the semi-structured interview, the researcher decided that piloting the questions with two participants was sufficient to ensure that the questions would elicit the data needed.

3.4.2 OPEN-ENDED QUESTIONNAIRE

3.4.2.1 Rationale

A questionnaire is used to collect large amounts of data from many respondents and if the questionnaire is returned anonymously, confidentiality of responses is ensured. In this study the questionnaire was considered to be the most effective and efficient instrument to collect data from the total population of teachers. It is particularly appropriate for this group of individuals since they are all part-time teachers with different schedules at different locations. Similar to interviews, questionnaires may have different forms and include either open-ended or close-ended items and sometimes both. An open-ended questionnaire was designed to elicit data in order to establish the construct of language proficiency. This data obtained from the EFL teachers were used to triangulate the data obtained from the learners.

3.4.2.2 Design of the Open-Ended Questionnaire

The teachers were asked the same questions as those of the semi-structured interview (see Appendix 2). They were asked their opinions about learners'

motivation for studying English, the contexts where learners would use the language and learners' expectations in terms of desired terminal language objectives. The Teachers' Questionnaire (TQ) was designed based on the SIP (after it was piloted) and consisted of the following three open-ended questions (see Appendix 3):

- Why do you think Egyptian adults study English? What are the different reasons that bring them to CACE?
- What are the different contexts or situations where your students would use English?
- If we look at English proficiency as a continuum that is divided into several stages starting with no knowledge of English until the "top", please describe the language of a person at the "top" of the English language proficiency continuum.

3.4.3 THINK ALOUD VERBAL PROTOCOLS

3.4.3.1 Rationale

A verbal protocol is data collected from an individual where the person is asked to 'think aloud' as he/she works through a task. The protocol consists of the individual's verbal report and the set of protocols gathered constitutes "a body of qualitative data" (Green, 1998). Verbal protocol analysis is a qualitative method particularly suited to investigating cognitive activity where "inferences are actually made about the cognitive processes that produced the verbalization" (Green, 1998: 1). Verbal reports are now being used more frequently in language test validation studies (Cohen, 1984 & 1994; Nevo, 1989; Alderson, 1990; Buck, 1992; Li, 1992; Cohen, 1998; Weir et al, 2000; Lazaraton, 2002) where students are required to verbalize their thoughts while they complete a test or a series of test items. Since this study is concerned with assessment validation and the focus is on identifying the test-taking strategies and the affective factors that affect performance, therefore, verbal reports are highly appropriate. As Cohen (2000: 128-129) states, "the reason why verbal reports have gained popularity in the last several decades, despite frequent criticism ... is that this research methodology provides data on cognitive processes and learner responses that otherwise would have to be investigated only indirectly." The use of verbal reports is unique in this particular context since there

only one published report of research conducted in Egypt to date (based on the data available to the researcher) using this method in language testing (Khalifa, 1997).

The two main categories of verbal reports are introspective or concurrent and retrospective. The distinction is based on the difference in time between processing and reporting. Concurrent reports are generated while the task is being carried out, while retrospective reports are generated after the task has been completed (Banerjee & Luoma, 1997; Green, 1998). There are difficulties with retrospective verbal reports since it involves recovering information from memory which makes them “susceptible to influences from unwanted variables” (Green, 1998:6). The informant may omit some information or include redundant information. Another problem is that the verbal “report may be contaminated by a subject’s efforts to ‘tidy up’ what happened or to rationalise what occurred” (p.10). Therefore, it is better to use concurrent verbal reports whenever possible since they are far less susceptible to these unwanted influences. In this study concurrent verbal reports were collected and recorded on tape.

Green (1998) describes a further classification of verbal reports and distinguishes between talk alouds and think alouds. In talk alouds, the report includes verbal information which are “the words in the mind or thoughts that might be spoken” (p.5). In think alouds, the report will include the verbal information plus non-verbal information. This distinction is difficult to maintain and informants producing verbal reports often are unable to differentiate between the two. Thus, it is recommended that think alouds are used, especially for tasks that require informants to report non-verbal information. In this study, think alouds are used since subjects are required to report their cognitive processes and feelings while taking a test.

Considering the validity of verbal reports, it is impossible to prove that verbal reports reflect the actual cognitive processes and feelings of the test-takers. However Ericsson and Simon (1993) in their extensive review of a large number of verbal protocol studies indicated that verbal protocol analysis is a valid and viable method provided that appropriate procedures are followed. Green (1998) stated that validity can be maximized by ensuring that clear and specific instructions are provided to respondents completing the verbal reports; discouraging respondents from “trying to explain or rationalize their thoughts”(p.11); and minimizing the intervention of the researcher while respondents are completing the verbal reports. In this study, the

researcher followed documented, clear and specific procedures for collecting the verbal report data.

Concurrent verbal reports have been criticized in that they may intervene and alter test-takers behavior and performance. Undoubtedly, the introduction of verbal reporting does alter the nature of the task and the test-taking task is no longer truly authentic. As Cohen (2000:143) states it is a “challenge for researchers ... to simulate task conditions as they would be if the instrument were administered without the addition of verbal reports” and to ensure that respondents complete the tasks as far as possible as they would without the verbal reports. Furthermore, in this part of the study the verbal reports were not used to examine test performance as such but the focus was on test-taking strategies and emotions expressed.

Another concern expressed is that individuals vary considerably in the quality and length of verbal reports produced. Green (1998) states that while there are individual differences in the quality and quantity of verbal reports produced, however, individuals tend to be consistent which allows us to conclude that VPA is reliable. She added that differences in verbal report data are attributed to task, contextual and individual difference variables rather than the lack of reliability of the VPA method. She recommended that sparse reports which represent incomplete records of thoughts being verbalized, are discarded.

3.4.3.2 Design of the Think Aloud Protocol

During this data collection phase, a new EPT (English language Placement Test) was being piloted in CACE. This provided the researcher with the opportunity to collect think aloud data from those who agreed to participate in the pilot study. Students were motivated to sit for the EPT by offering them a certificate indicating their level of proficiency as well as the opportunity to enroll in higher proficiency level classes should their results on the pilot EPT prove higher than their current level.

The respondents were not required to explain their thoughts which may have changed the way they approach the task. After giving the instructions, the researcher sat at a distance from the respondent and only intervened, by instructing or reminding the respondent to keep thinking aloud, when the test-taker fell silent for a period of time. Therefore, prompts were only used when a test-taker was silent for a while and the researcher would prompt the test-taker by stating “please continue to

think aloud” or “please don’t stop talking”. In order to collect adequate data, respondents were encouraged to think aloud in Arabic or in both English and Arabic as appropriate.

3.4.3.3 Piloting of the Think Aloud Protocol

The instructions used were those adapted from Green (1998:46) as follows:

I am interested in what you think about as you complete the test I am going to give you. To do this, I am going to ask you to think aloud as you work through the test. By ‘think aloud’ I mean that I want you to say out loud everything that you are thinking and feeling from the time you start the test until the end. I would like you to talk constantly from the time you begin the test until you have finished it. You can talk in Arabic only or in both English and Arabic. Do not plan or try to explain to me what you are thinking. It may help if you imagine that you are in the room by yourself. It is very important that you keep talking. If you are silent for any period of time, I will remind you to keep talking.

Do you understand what I am asking you to do? Do you have any questions?

We will start with a practice question. First I would like you to think aloud as you complete the following question:

Hurry! The train _____ in five minutes.

- (A) left
- (B) leaving
- (C) has left
- (D) is leaving

The instructions were given in Arabic to all respondents. These instructions were piloted on a respondent and it was found that she had difficulty in producing a verbal report. The instructions were then modified by the researcher first demonstrating a think aloud on a test question and then getting the respondent to complete the think aloud on the practice question as follows:

I am interested in what you think about as you complete the test I am going to give you. To do this, I am going to ask you to think aloud as you work through the test. By 'think aloud' I mean that I want you to say out loud everything that you are thinking and feeling from the time you start the test until the end. I would like you to talk constantly from the time you begin the test until you have finished it. You can talk in Arabic only or in both English and Arabic. Do not plan or try to explain to me what you are thinking. It may help if you imagine that you are in the room by yourself. It is very important that you keep talking. If you are silent for any period of time, I will remind you to keep talking.

Do you understand what I am asking you to do? Do you have any questions?

I will start by demonstrating for you a think aloud on a test question.

You _____ eat any kind of food that has reached its expiry date or you may die of food poisoning.

- (A) don't have to
- (B) must not
- (C) might not
- (D) ought not

Sample concurrent verbal report (demonstrated by the test administrator or researcher)

Well let me read. The question looks very long. There are many difficult words. I don't understand expiry. I don't know the meaning of poisoning. I have to guess. I feel confused and frustrated. It's too difficult for me. I will choose answer 'A' because it looks different from the other three. The three all have "not" while 'A' is different.

Now let's complete a practice question. I would like you to think aloud as you complete the following question:

Hurry! The train _____ in five minutes.

- (A) left
- (B) leaving
- (C) has left
- (D) is leaving

These were the final instructions used to collect the verbal report data in this study and they were given in Arabic.

3.4.4 LIKERT-TYPE QUESTIONNAIRE

3.4.4.1 Rationale

In order to further explore the findings of the data from the think alouds, a qualitative data collection method involving a limited number of informants, and to determine the relationship between these findings and performance on a test, a quantitative survey of a sample of this population, was conducted. The Likert-type questionnaire was considered most suitable since it is an effective means of collecting a large number of responses. The questionnaire was designed based on the findings from the think aloud verbal protocols and based on test-taking strategies identified from the literature.

3.4.4.2 Design and Piloting of the Test-taking Strategies Questionnaire (TTSQ)

The TTSQ (see Appendix 4) consisted of two sections and included both types of items: open (yielding qualitative data) and closed (yielding quantitative data). Section 1 focused on test-taking strategies used while taking a placement test and included forty-nine (49) closed items and one open-ended item. In the closed items, respondents were required to read and rate each statement in accordance with the following 3-point scale: 3 = I used this strategy several times during the test; 2 = I used this strategy only once or twice during the test; 1 = I did not use this strategy during the test. The 3-point scale was used to identify the strategies that were and were not used and to differentiate between high frequency and low frequency strategies. Thus, a rating of 3 signified a test-taking strategy that is frequently used, a rating of 2 indicated a low frequency test-taking strategy and a rating of 1 meant that the strategy was not used. The test-taking strategies were grouped into four categories: general strategies not specific to a section or language skill, strategies used in the writing section, strategies used for listening and strategies used in the reading comprehension section. The strategies were not listed in any specific order. The open-ended question focused on possible test-taking strategies not mentioned in the questionnaire. If respondents stated that they used a strategy not mentioned, they were required to describe the strategy and indicate whether they used it once or twice or more.

The forty-nine test-taking strategies were obtained from both the literature and the think aloud data reported on in detail in Chapter 5. Thirty-two (32) of the forty-nine (49) strategies in the TTSQ were obtained from the literature. Table (5)

Table (5): List of test-taking strategies obtained from the literature* and included in the TTSQ

Strategies for Taking a Multiple Choice Proficiency Test Not Specific to One Skill

1. Reading the questions and options before choosing one
2. Stopped reading options when they got to the one that seemed correct.
3. Selecting an option by eliminating the other 3 options
4. Making an educated guess using background knowledge
5. Trying to produce their own answer to the question before looking at the options provided
6. Rereading the questions and options for clarification
7. Postponing dealing with a question or selecting a given option until later.
8. Skipping a question that is not understood and leaving the answer blank
9. Guessing without any particular considerations
10. Changing responses when appropriate
11. Selecting an option that is longer / shorter than the others
12. Looking for an option that seems to be different from the others
13. Running out of time without trying all the questions
14. Trying to finish the test as fast as possible
15. Previewing or surveying the whole test
16. Monitoring time
17. Watching to see when other students finish the test.
18. Reading the instructions carefully.
19. Never leaving an answer blank
20. Going back and review or check answers.
21. Getting stuck on one question for a long time.

Specific Strategies for Reading

22. Reading the passage first
23. Reading the questions first before reading the passage.
24. Translating relevant parts of the passage for understanding.
25. Guessing meaning of an unknown word from the context
26. Skipping unknown words.
27. Using knowledge of grammar to answer the question.
28. Looking for portion of the passage that the question refers to and then look there for clues to the answer
29. Matching material from the passage with material in the question and in the options
30. Selecting an option because it appears to have a word or phrase from the passage in it
31. Selecting an option based on understanding the passage read
32. Getting clues from answering one question that were helpful in answering another question

* The sources in the literature drawn upon are listed before the table

shows the list of test-taking strategies included in the TTSQ found in the literature. The sources of these strategies in the literature are: Cohen's (1998) list of Strategies for taking a Multiple Choice Reading Comprehension Test, Anderson's (1991) Categorization of Processing Strategies and test-taking strategies were identified from Vattanapath & Jaiprayoon (1999) and Roth et al (2000).

Seventeen (17) test-taking strategies that were reported by the respondents in this study during the think alouds and which were not included in any of the lists found in the literature were added to the list of strategies. The strategies obtained from the think aloud data are reported in section 3.10.2.2.

Section 2 focused on the affective factors of taking a test i.e. the feelings that occurred during the test. This section consisted of twenty-nine (29) closed items and three open-ended items. In the closed items, respondents were required to read and rate each statement in accordance with the following 6-point scale: 6 = I strongly agree; 5 = I agree; 4 = I slightly agree; 3 = I slightly disagree; 2 = I disagree; 1 = I strongly disagree. The 6-point scale was used to eliminate neutral responses. A 6-point scale was used in favor of a 4-point scale because increasing the number of steps or points in a scale increases its reliability:

"The number of steps in the scale is important. The more steps, the greater the reliability of the scale. This increase is noticeable up to 7 steps, the increase in reliability declines as 20 steps are reached. Items with more than 7 steps are seldom used because the increase in reliability resulting from additional steps is slight. It is easier to increase reliability by adding more items."

(Thorndike et al, 1991:314).

The first open-ended question requested respondents to describe their feelings during the test. The second question focused on whether respondents experienced anxiety in a specific section or sections of the test and the third question required respondents to indicate whether anxiety differed from one section of the test to another. The reason for including these questions was that the think aloud verbal protocol data showed that some respondents experienced far more anxiety in the listening section compared to other sections of the test.

The twenty-nine closed items were obtained from both the think aloud verbal protocol data and the literature. All the emotions, emotional regulation processes and behaviors obtained from the think aloud data were found in the literature except for two items related to having the instructions and writing prompt in both Arabic and English (shown in italics in Table (6)). There were eight (8) items obtained from the literature (Bradshaw, 1990; Vattanapath & Jaiprayoon (1999) and Roth et al (2000)) only and not from the think aloud verbal protocol data which are shown in bold in Table (6).

The items were not listed in any particular order. Table (6) shows the items classified by emotions, emotional regulation processes and behaviors.

Table (6): Emotions, emotional regulation processes and behaviors obtained from the think aloud data and the literature

TTSQ Item	Emotion / Emotional Regulation Process
Tried hard on the test	Behavior
Did well	Assessment
Test or item was easy	Assessment
Was so anxious felt like getting the answer from another person	Emotion
Test / item was confusing	Emotion
Test / item was difficult	Assessment
Felt prepared	Emotion
Knows what to do	Task focusing
Thought it was important to do their best	Assessment (goal orientation)
Had enough time to finish the test	Task focusing
Felt that taking the test was a big challenge	Emotion
Felt a sense of achievement after completing a section, or an item or the test.	Emotion
Felt relieved when the test was over	Emotion
Felt nervous during the test	Emotion
Had difficulty in concentrating	Emotion
Kept looking around the test room during the test	Behavior
Dreads taking tests in general	Emotion
Dreads taking tests because they don't show true ability	Emotion
Felt tired during the test	Emotion
Gave up because the test was too difficult	Behavior
Got tired and started answering without reading the question	Emotion
Because of nervousness forgot the things that they usually know	Emotion
Was sure of the correct answer	Task focusing
Felt the test was interesting	Emotion
Felt bored while taking the test	Emotion
Taking the test was a pleasant experience	Emotion
Felt frustrated because there was not enough time.	Emotion
<i>Thought it was a good idea to have the instructions in both Arabic and English. (task focusing)</i>	Task focusing
<i>Thought it was a good idea to have the writing prompt in both Arabic and English. (task focusing)</i>	Task focusing

The feeling that taking a test was a big challenge is described as the feeling that the test is rather daunting or overwhelming and thus, it is associated with a negative emotion.

The questionnaire was administered in Arabic. A preliminary version of the questionnaire was formulated in English and then translated into Arabic. The preliminary version (see Appendix 6) was based on the list of codes (see section 3.10.2.1) used to analyze the think aloud data with some changes as follows:

- Added: kept looking around the testing room during the test
- Deleted: felt surprised at having the writing prompt both in Arabic and English
- Deleted: disapproved about having the writing prompt both in Arabic and English

The two items were deleted because only one test-taker mentioned each of these feelings and they are redundant. The preliminary version of the questionnaire consisted of 77 items.

The questionnaire was then piloted on two adult EFL learners: one female and one male. The researcher sat with each respondent separately and requested him / her to read the directions, the descriptions of the rating scale and all the items and explain to the researcher what he / she understood. Accordingly, based on the feedback received from the two respondents the following items were changed:

- “I skipped a question that I did not understand and left the answer blank.” was changed to “If there was a question that I did not understand, I left the answer blank.”
- “I selected an option that was longer / shorter than the others.” was changed to “When I was not sure of the answer, I selected an option that was longer / shorter than the others.”
- “I looked for an option that seemed to be different from the others.” was changed to “When I was not sure of the answer, I looked for an option that seemed to be different from the others.”
- “I read the options of the next question before listening to the question on tape.” was changed to “I had sufficient time to read the options of the next question before listening to the question on tape.”

- “I felt like cheating.” was changed to “I felt so anxious that I wanted to get the answers from another person.” Both respondents thought that the word cheating has such negative connotations and would not lead to direct answers.
- The item “I enjoyed taking the test” was deleted because they felt it was redundant.
- Both respondents found the item “I thought taking the test was a challenge.” to be confusing since it could be interpreted either positively or negatively. The item was rewritten as two new items: “I felt that taking the test was a big challenge.” and “I felt a sense of achievement after completing the test.”
- Both respondents found the item “I dread taking tests because I know more than the test will show.” to be confusing. The item was rewritten as two new items: “I dread taking tests in general.” and “I dread taking tests because I know that they do not show my true ability.”

Based on a suggestion from one of the respondents, the format of the questionnaire was changed to a table format which made it easier for test-takers to complete. A similar suggestion was made for displaying the rating scales. While the descriptors of the two rating scales were clear, the scales were also presented in a table format in the final version of the questionnaire.

In order to ensure that all the items were written and phrased correctly in Arabic, the Arabic version of the questionnaire was then back-translated into English. As a result of this back-translation the Arabic version of the following items were modified to ensure that they matched the intended meaning as they were originally formulated in English:

Section 1:

- | | |
|----------|--|
| Item 6: | I made an educated guess using background knowledge. |
| Item 11: | I guessed without any particular considerations. |
| Item 28: | I reread the prompt. |
| Item 30: | I used words from the prompt in my answer. |
| Item 39: | I translated relevant parts of the passage to understand. |
| Item 41: | I was able to guess the meaning of an unknown word from the context. |
| Item 42: | I skipped unknown words. |

Item 47: I selected an option because it appeared to have a word or phrase from the passage in it.

Section 2:

Item 5: I felt the test was confusing.

Item 9: I thought I knew why it was important to do my best.

Item 22: I forgot the things I usually know because I was so nervous.

Item 29: I think it is a good idea to have the writing prompt in both Arabic and English.

The Arabic version of the questionnaire and the English back-translation are shown in Appendix 4.

3.5 RESPONDENTS

The population of this study consisted of Egyptian adult EFL learners and EFL teachers in the English Studies Division of CACE, AUC (The Center for Adult & Continuing Education at the American University in Cairo). When collecting data using the semi-structured interviews, open-ended questionnaire and the think aloud protocols, the English language program consisted of six proficiency stages. In 2000 the program was restructured and the new curriculum consisted of four proficiency stages. Therefore, when collecting the TTSQ data, the respondents were subsequently placed into these four stages: novice, elementary, intermediate and advanced. There were four groups of respondents in this thesis: a different group for each of the four data collection instruments used.

3.5.1 Respondents in the Semi-Structured Interviews

The respondents were Egyptian adults applying to enroll in an English language course in CACE, AUC. A randomly selected sample of thirty-six (36) students were interviewed. The students selected were representative of the different groups of the population based on the following three background variables: proficiency level: elementary, intermediate and advanced (the six stages that were in use in the institution were grouped into three proficiency levels), gender (male, female) and age (two age groups, group 1: 18-22, group 2: 23+). These background variables were identified as significant in previous studies of the same population. The six stages at CACE were: basic, elementary, lower intermediate, intermediate,

upper intermediate and advanced. Basic and elementary were combined into elementary, lower intermediate and intermediate into intermediate and upper intermediate and advanced were grouped as advanced. Regarding age, the subjects were divided into two main groups: those who were still studying and not working and those who were working. The following table shows the distribution of the respondents:

Table (7): Semi-structured interview respondents

AGE	ELEMENTARY		INTERMEDIATE		ADVANCED	
	Male	Female	Male	Female	Male	Female
18 – 22	3	3	3	3	3	3
23+	3	3	3	3	3	3
	6	6	6	6	6	6
Total	12		12		12	

Random sampling was used as a strategy to ensure that various groups of the population are represented. However, within this research design, random sampling does not guarantee generalizability as discussed in section 3.7.

3.5.2 Respondents to the Open-Ended Questionnaire

There was a total of sixty-seven (67) EFL teachers in the program and they were all surveyed using a questionnaire. Forty-one (41) teachers completed the questionnaires, representing a response rate of about 60%.

3.5.3 Respondents in the Think Aloud Verbal Protocols

Twelve (12) test-takers were asked to ‘think aloud’ and to verbalize their feelings as they took the test described above. Test-takers verbal reports were tape-recorded. The twelve respondents are representative of various groups in the population based on three background variables: gender (male, female), age (18-22, 23+) and English language proficiency (elementary, intermediate, advanced). In order to obtain data from twelve respondents, data were obtained from seventeen test-takers. Five tapes that were not clear or did not contain sufficient introspective data were discarded. Some test-takers were not able to think aloud or were reluctant to do so after they started although they had initially agreed to participate in the study. The five respondents whose tapes were discarded were: three males at the elementary stage & age (23+), one female at the intermediate stage & age (18-22)

and one female at the advanced stage & age (23+). The following table shows the distribution of the twelve (12) respondents:

Table (8): Respondents in the think alouds

AGE	ELEMENTARY		INTERMEDIATE		ADVANCED	
	Male	Female	Male	Female	Male	Female
18 – 22	1 (A)	1 (B)	1 (E)	1 (F)	1 (I)	1 (J)
23+	1 (C)	1 (D)	1 (G)	1 (H)	1 (K)	1 (L)
	2	2	2	2	2	2
Total	4		4		4	

The twelve test-takers are referred to by letters: A – L. The selection of respondents is described in section 3.6.3.

3.5.4 Respondents to the TTSQ (Likert-Type Questionnaire)

The respondents were Egyptian adults applying to enroll in an English language course in CACE, AUC. A total of 538 respondents completed the questionnaire. Questionnaires from respondents who did not complete at least 80% of the items were discarded resulting in a total of 497 questionnaires used for the final analysis.

3.6 DATA COLLECTION PROCEDURES

3.6.1 Data Collection Procedures for Semi-Structured Interviews

With the cooperation of the admissions department, students with the required profiles were invited to meet with the researcher. In compliance with the ethical guidelines for conducting research, the researcher explained the purpose of the study i.e. investigating learners’ perceptions of English language proficiency, and obtained their agreement to participate and to tape-record the interview. Each respondent was informed that the interview would last approximately fifteen (15) minutes.

Most of the interviews were conducted in colloquial Arabic. Arabic is a diglossic language where there are two distinct varieties of the language, colloquial Arabic and classical Arabic. Colloquial Arabic is the normal spoken language and classical Arabic is used for written purposes. Since colloquial Arabic is not a written language, all of the taped interviews were directly translated from colloquial Arabic

and documented in English. A record of the responses to the questions for each interview were all transcribed on to disk and all the data are shown in Appendix 7 demonstrating transparency and ethics in conducting research.

3.6.2 Data Collection Procedures for the Open-Ended Questionnaire

The researcher obtained the general schedule of all the teachers and distributed the questionnaire to each teacher by hand. An envelope was also provided which included the researcher's university internal mail address. Teachers were requested to return the questionnaires by hand or internal mail. To ensure confidentiality and to conform to ethical guidelines for conducting research, teachers were not required to write their names on the questionnaire. All the data are shown in Appendix 8 to demonstrate ethical practice in this research study.

3.6.3 Data Collection Procedures for the Think Aloud Verbal Protocols

In order to select respondents who would volunteer for this type of task, the researcher required the assistance of teachers in this process. One teacher from each stage of the English language program at CACE: elementary, intermediate and advanced, was randomly selected and contacted by the researcher. The researcher met with each teacher, explained the purpose of the study and what was required and obtained their agreement to participate in the study. Each of the three teachers was requested to select eight students from the class(es) they teach matching the required profile in terms of gender and age who would volunteer to participate in the study (the researcher needed only twelve test-takers but requested twenty-four students in order to guarantee a final sample of twelve). It was emphasized to the three teachers that they should select students with eight different profiles. This meant that teachers had to find out the age of the students they selected. Teachers were requested to explain that the new CACE placement test was being piloted and that the purpose of this study was to investigate students' thoughts and feelings as they sat for the test. Teachers were requested to stress to those who volunteered that the result would not affect their current placement at all. In order to ensure that the goal for taking the test would be important and relevant to the test-takers, the students were informed that they would be provided with a certificate of their test results. In the Egyptian context, a certificate from CACE, AUC indicating their proficiency level is a very strong motivator for many students. Furthermore, they were told that should their

results on the pilot EPT indicate they should be placed in a higher level, they could take advantage of this if they desired in the following term. The researcher emphasized that it should be made clear to the students that they were being asked to volunteer about two hours of their time to participate in the study.

Teachers were also asked to arrange for a suitable time at the convenience of the student to take the test and to explain that taking the test did not have any negative consequences for the student. The researcher provided each teacher with a schedule of a range of possible time slots (in the morning and evening) that students could choose from. All the think alouds were conducted at the Educational Assessment Unit of CACE and the students were given clear directions by the teachers. The teachers sent the researcher a list of names of students with their profiles and the timetable.

Out of the twenty-four students who initially volunteered, seventeen students actually participated in the study (seven students did not show up). The researcher selected the final sample of twelve think aloud protocols from the seventeen by examining the quality of the think aloud data (as described in section 3.5.3) and ensuring that the three background variables of the population were represented in the final sample. In case where two protocols were similar, the researcher randomly selected one of them to include in the final sample.

In this Think Aloud study Form A1 of the English Proficiency Test (EPT) was used to collect the verbal protocol data (see Appendix 9 which includes the instructions and example items. The actual items are not included for test security since the test is still in current use). The time limit for sections two and three was prolonged for the verbal report. Five additional minutes were added to the grammar and reading comprehension sections. It was decided that giving test-takers more time may contaminate the actual testing situation and thus, it was limited to only five extra minutes for each section. It was not possible to add time to the listening comprehension section, however, in the design of the EPT, the time allowed between each item ranged from 18 – 20 seconds which is considerably higher than the TOEFL or TOEIC test where the time interval is 12 seconds.

To ensure conformance with ethical guidelines for conducting research all students were assured of confidentiality of any information obtained. All of the taped think alouds were directly translated from colloquial Arabic and documented in English. A record of the responses were all transcribed on to disk and four

samples are shown in Appendix 10. Appendix 11 shows four samples of the correct / incorrect responses for each multiple-choice item for each test-taker and a summary of the comments made for each question. These samples are included to demonstrate conformance to ethical guidelines for conducting research.

3.6.3.1 Description of the Test Task

The EPT is the test used to place students in the English language program of the Center for Adult & Continuing Education. The EPT has two components: a writing component and a multiple-choice questions (MCQ) component. In the writing component test-takers are required to write a business letter in 30 minutes. Test-takers are given a prompt that describes a particular work-place situation where a letter should be written. The prompt is also translated into Arabic to ensure that test-takers understand the situation.

The EPT multiple-choice component consists of three sections:

- Section 1: Listening Comprehension (20 minutes)
- Section 2: Grammar (20 minutes)
- Section 3: Reading Comprehension (30 minutes)

The Listening Comprehension section consists of three parts and is designed to measure test-takers ability to understand spoken English. In Part A, test-takers listen to fifteen (15) statements which are spoken once only. After each statement, test-takers are required to select out of a possible four pictures the one picture which corresponds to the meaning of the statement they had just heard. In Part B, test-takers listen to fifteen (15) questions and are required to select the best answer to the question they listened to. In Part C, test-takers listen to ten mini-dialogues between two speakers. At the end of each mini-dialogue, a third person asks a question about what was said and test-takers are required to select the best answer to the question.

The Grammar section consists of two parts and is designed to measure test-takers ability to recognize language that is appropriate for standard English. In Part A, there are fifteen (15) incomplete sentences or questions and test-takers are required to select the word or phrase that best completes each item. In Part B, there are fifteen (15) sentences or questions with one error and test-takers are required to identify this error. Each sentence or question has four underlined words or phrases and test-takers choose the one word or phrase that must be changed in order for the sentence or question to be correct.

The Reading Comprehension section consists of a variety of reading texts followed by three, four or five multiple-choice questions on each text resulting in a total of twenty (20) questions. There are five reading passages which include a memo, a letter, a short report and two general texts.

The format of the test, weighting and scores assigned to each test section are as follows:

<u>Testing section</u>	<u>No. of items</u>	<u>Allocated time</u>	<u>Scores</u>
Writing	One	30 minutes	25 points
Listening	40 MCQ	approx. 25 minutes	40 points
Grammar	30 MCQ	20 minutes	30 points
Reading	20 MCQ	30 minutes	20 points

The total number of points of the MCQ sections (90 points) is converted to a score out of 75 points and then added to the score of the Writing section (25 points) to obtain a total score out of 100. In a validation study conducted by CACE, the cut-offs were established for the four proficiency stages as follows: novice (0 – 22%), elementary (23 – 54%), intermediate (55 – 86%) and advanced (87 – 100%).

The internal consistency reliability of the MCQ component of the test is very high. The Kuder-Richardson 20 (KR-20) coefficient is 0.95.

3.6.4 Data Collection Procedures for Administering the TTSQ

The questionnaire was administered right after the respondents completed the English Proficiency Test. The questionnaire required approximately 15 – 20 minutes to complete. The data were collected at the end of four testing sessions. At the beginning of each testing session, the test administrator made the following announcement to the test-takers:

“The testing specialists at the American University in Cairo are conducting a survey in order to find out the different strategies used and the reactions to the test taken by CACE applicants. The objective of this survey is to improve the placement testing services offered by CACE and is part of the university’s overall plan for continuous improvement. Because you are a representative group of CACE applicants, we kindly request you to stay after the end of the test and complete a questionnaire for us.”

At the end of the test after all the test papers were collected, the test administrator requested all the test-takers who wanted to complete the questionnaire

to remain in their seats. Ensuring ethical practice in conducting research, any test-taker who had other commitments and was not able to stay to complete the questionnaire was allowed to leave. Test-takers whose native language was not Arabic were asked to leave since the questionnaire was administered in Arabic. The questionnaires were distributed and the following instructions were given:

“Please concentrate while reading the items and choose your responses as honestly and carefully as possible because the information obtained will have a significant impact in the design of future tests. It is important to note that there are no right or wrong answers. Your responses to this questionnaire have no effect on your test result whatsoever and will be completely confidential and accessible only to the researcher conducting this study. Because the aim of this study is to investigate the relationship between test-taking strategies used and test performance, you are kindly requested to write down your student number and name.”

The test administrator then proceeded to explain the two rating scales. He read aloud each descriptor and then proceeded to demonstrate how each scale works on item 1 in Section 1 and item 1 in Section 2. Respondents were then asked to complete the questionnaire and indicate when they finished by raising their hands. The test proctors circulated around the testing room collecting the completed questionnaires.

The test papers were scored and the researcher obtained an Excel file with the test-takers' results and placement in one of the four proficiency levels of CACE. The proficiency levels are: novice, elementary, intermediate and advanced. The researcher then matched the test-takers' test results and placement levels with the completed questionnaires. This matching process was based on the student number and names on the test results and questionnaires.

3.7 GENERALIZABILITY AND RELIABILITY

Generalizability refers to the extent that results obtained from samples of individuals or events are generalizable to the population. Bassey (1990:19) states that “generalisation is a statement which collates evidence of particular events, and extrapolates that evidence to predict the occurrence of similar events.” From the quantitative perspective, representative samples are required for generalizability and researchers use probability sampling strategies. The number of observations is important for quantitative researchers to ensure generalizability since “sample size

affects the margin of error and the power of statistical tests to detect effects” (Tashakkori & Teddlie, 1998: 72). For a case study which comes under qualitative methods, generalizability of findings to other individuals or contexts is not desired and is not the focus. Qualitative researchers use purposive sampling since the selection of individuals is based on the specific research questions being investigated rather than on representativeness. Some qualitative researchers refer to generalizability as transferability. However, despite these distinctions some degree of generalizability or transferability is important to all researchers (Tashakkori & Teddlie, 1998). Generalizability of case study findings to other situations is difficult, however researchers can use strategies to help the ‘consumers’ of the case study data determine how applicable the results are to their particular situation. Bassey (1990) discussed the concept of relatability where the consumers of a research study can relate the findings to their own situation. Although situations are different, often there are similarities that can be related to each other. Thus, in order to achieve transferability or relatability of the interpretations and inferences made, it is important to include sufficient details of the participants and context that comprise the case such that the inferences can be transferred to another context (Gall et al, 1996). A detailed description of all relevant information is provided throughout this thesis. Thus, a key strategy used to ensure external validity is the provision of a detailed description of all information (i.e. thick descriptions) about the context in order to provide evidence for transferability of inferences (Creswell, 1994).

Random sampling was used in the administration of the semi-structured interviews and the Likert-type questionnaire. The Likert-type questionnaire was administered to a total of 497 students (a sufficiently large sample from which generalizations can be made to the population) and the Teacher Questionnaire was administered to the total population of EFL teachers in the program. For the verbal protocol analysis, 12 respondents were selected representative of the various groups of the population as identified by the researcher. Purposive sampling was used since subjects had to volunteer and be trained for this type of task: to think aloud while taking a test.

According to Tashakkori & Teddlie (1998) when a mixed methods design is used to examine the various research questions, there is switching between the different modes of generalizability. Quantitative methods focus on large random samples of respondents and qualitative research seeks answers to how respondents

make sense of their experiences and therefore, qualitative sampling tends to be purposive. Although, this research study used a mixed methods design, however, the researcher does not agree that there are different modes of generalizability in this study. As mentioned, this study is a case study conducted in a specific context and it is not possible to generalize. Although random sampling and large samples were used in various stages of data collection, however, it does not ensure generalizability.

This study was conducted at CACE, AUC and therefore, the findings cannot be generalized to all Egyptian adult EFL learners. AUC is an American institution and learners with particular profiles may choose to come to study English for this reason while other Egyptian adult EFL learners with different profiles may choose to study English at the British Council for example. The context of this study is located in Cairo, the largest city in Egypt, which is very different from all other cities and in particular from the south of Egypt where it is more rural and less developed.

Thus, in order to ensure an adequate degree of transferability or relatability of findings from this study to other contexts, the researcher has provided detailed descriptions of the contexts and participants in this study.

3.8 VALIDITY

Validity is an indication of how sound the research is and it applies to both the design and methods used. Validity is defined as "how accurately the account represents participants' realities of the social phenomena and is credible to them" (Creswell & Miller, 2000:124). Research is viewed in terms of external and internal validity. External validity is defined as the extent to which the research findings are generalizable from a sample to a population or from one setting to other similar settings or the extent to which the findings are transferable or relatable from one context to another, and this has been discussed in the previous section.

Internal validity is defined as the extent to which we can trust the research findings or inferences regarding the relationship between variables. If the obtained relationship between variables is real and not based on alternative explanations then the conclusions regarding the relationship between variables have internal validity. In qualitative research, internal validity is viewed as "how far the researcher's constructions are grounded in the constructions of those whom he or she studied and how far this grounding is transparent for others" (Flick, 2002:222). Therefore, the methods used to produce the data is the starting point for judging their validity and

the subsequent interpretations. To establish internal validity, the categories and interpretations of the researcher are presented in detail and examined to determine whether they reflect reality or whether the findings are distorted by limitations in data collection.

Seliger (1983) has argued that "self-report data cannot be independently confirmed and should, therefore, be treated with caution" (Purpura, 1999:11). Seliger (1983:183) stated that "introspections are conscious verbalizations of what we think we know" and "conscious verbal reports of learners about their own internal device cannot be taken as a direct representation of internal processing" (p.189). It is recognized that self-report data present a threat to internal validity and that data elicited from a questionnaire are not a direct reflection of mental processing. It is also recognized that respondents may misunderstand an item or may lack self-awareness. Respondents may select certain items on self-report instruments that they perceive as more socially desirable responses than other items (Allan, 1995). To compensate for these potential threats to internal validity, the use of multiple data collection methods is recommended and is used in this study. "Verbal reports that involve intervening during the performance of a task have been criticized for the reactive effects that such interventions may cause" (Cohen, 2000). In order to ensure that the verbal report does not interfere with the task, Ericsson & Simon (1993) recommend that respondents are provided with warm-up tasks to train them to make verbal reports without confounding them with explanations of their responses. In this study that researcher has provided clear directions and orientation to the respondents providing the think aloud data.

There are specific features in the design of this study that address internal validity. The following strategies are employed to ensure internal validity: triangulation, piloting of all the data collection tools used and clarification of the researcher's role. This study uses triangulation techniques by using a combination of quantitative and qualitative methods. Denzin (1978) identifies four types of triangulation: data, researcher, theory and methodological triangulation. Data triangulation involves the use of a variety of data sources, researcher triangulation refers to the use of several different researchers, theory triangulation is the use of more than one perspective to interpret findings and methodological triangulation involves the use of multiple methods. In this study both teachers and learners were a source of data and the data were collected across different time frames. Investigator

triangulation was not applied in this study. Regarding theory triangulation, it is difficult to mix the two paradigms: positivist and interpretivist in the use of theory since they differ in epistemology, axiology, ontology, causality and methods used. It is problematic to locate research issues in both the positivist and interpretivist paradigms. In this study, a pragmatic perspective is used in addressing the research issues (Allan, 1996; Tashakkori & Teddlie, 1998). In fact Denzin (1978:307) questioned the plausibility of theoretical triangulation:

"My use of theoretical triangulation must in no way be construed as a defense of eclecticism. Indeed, sociologists committed to a given perspective will probably not employ theoretical triangulation. The great value of this strategy, as I see it, however, is its assurance that no study will be conducted in the absence of some theoretical perspective. In this sense it is most appropriate for the theoretically uncommitted, as well as for analysis of areas characterized by high theoretical incoherence."

Multiple methods are used at different phases of the research in order to better understand a concept or construct being explored. Initially, the assumption underlying the use of triangulation was that bias inherent in any one method would be eliminated by using multiple methods. Triangulation was seen to provide evidence resulting in convergence on a single proposition about some social phenomenon. This view was challenged by Mathison (1988) and she stated that in actual practice triangulation often resulted in inconsistency or contradiction. It is rare for researchers to end up with data that converge on a single proposition. It is the task of the researcher to make sense of the evidence. Triangulation is seen as a "state of mind" and Mathison states that "the value of triangulation lies in providing evidence – whether convergent, inconsistent or contradictory – such that the researcher can construct good explanations of the social phenomena from which they arise" (p.15). Triangulation provides a richer and more complex picture about the social phenomenon being studied (Mathison, 1988).

Both data triangulation and multiple data collection methods were used. A wide range of views regarding English language proficiency were obtained from semi-structured interviews with Egyptian adults from a variety of backgrounds. Further data were obtained from the perspective of teachers. Thus, triangulation was used to identify potentially inconsistent or contradictory evidence. Triangulation

was also applied by using multiple methods sequentially. For example, think aloud protocols were used to explore the test-taking strategies used by Egyptian EFL learners and then a Likert-type questionnaire was used to identify the relationship between affect, test-taking strategies and performance on a test. Thus, it has been demonstrated that the emergent constructs related to language proficiency and test-taking strategies are soundly based on a combination of quantitative and qualitative methods.

The second strategy used to ensure internal validity was the piloting of all the data collection instruments. In order to minimize misinterpretation of the questions in the semi-structured interview and the questionnaires, all the instruments were piloted and tried out beforehand. This ensured that all the respondents understood the wording or the meaning of the questions in the same way. Furthermore, the procedure for collecting the think aloud data was piloted to ensure that the respondents understood what was required.

In order to strengthen the validity and precision of the results of the qualitative data and the inferences made researchers must be aware of the context, the culture and multiple perspectives of informants. Thus, the researcher's role and values and judgment of the researcher are very important in this study. The researcher is very familiar with the specific context having had taught and conducted several research studies in it for many years. Accordingly, the researcher has had adequate time in the field to "learn the culture, and test for misinformation either from informants or from their own biases" (Lincoln & Guba, 1985:90). Furthermore, it is important for the researchers to acknowledge and disclose their beliefs and biases early in the research process. At the beginning of this chapter, the researcher described in detail her paradigm position which reflects her personal beliefs and biases that shape this enquiry.

The construct definition in this thesis of English language proficiency is consistent with the findings of a subsequent study (Aydelott et al, 2000) conducted with a sample of 658 students and 74 teachers in the same program. This study provides further empirical evidence of the validity of the components of the construct being defined (more details are provided in Chapter 4). Having discussed the validity of the study, issues related to reliability will be discussed in the following section.

3.9 RELIABILITY

3.9.1 Procedural Reliability

Reliability is defined as the stability of measurements or observations. Reliability also includes the quality of recording and documenting data. In order to increase the reliability of data, standardization of procedures for collecting data is required. In this study detailed protocols for collecting data at different phases of the study are documented and reported. Furthermore, all the interview and the think aloud data were tape recorded and then transcribed. Thus, the data was made explicit. This detailed documentation of the procedures for collecting and obtaining explicit data is the technique used to ensure the dependability or reliability of the results and interpretations (Flick, 2002).

3.9.1.1 Reliability of the Think Aloud Verbal Reports

Pressley & Afflerbach (1995) suggest nine areas that researchers should focus on when reporting on verbal report data that would ensure comparison across studies. Characteristics of respondents such as their educational background, knowledge of the task, motivation to do the task, level of language proficiency, and age should be clearly reported. Researchers should also indicate the language of the verbal reports produced. In this study, the researcher provided all the details on the characteristics mentioned except educational background since it was not deemed relevant.

The characteristics of the materials used in the task and the task itself should be described and this was done so. The researcher described the test used in the think aloud in detail and it was demonstrated that the test was designed based on a construct that was specific to the context of this study. The nature and extent of guidance that was provided during the think aloud is important for comparing research studies and details were discussed in section 3.6.3 of this chapter. The methods of analysis and the categories used to score the verbal report protocols are also described in the following section 3.10.2 of this chapter. Inter-rater reliability checks are reported in the following section and samples of the transcripts of the verbal reports produced are included in Appendix 10. Finally, the theoretical principles that the verbal reports were investigating were clearly described and discussed in detail in the literature review (Chapter 2). Therefore, it has been

demonstrated that the verbal report procedures were rigorous in design and description.

3.9.2 Reliability of the TTSQ

The internal consistency reliability of the Likert-type questionnaire was assessed by means of Cronbach's alpha coefficient. The Cronbach alpha coefficient is the most commonly used indicator of internal consistency in conjunction with Likert-type scale items. Each item is compared with each other item to check for consistency of the rating scale. The overall reliability of the TTSQ is 0.78. The reliability of Section 1 of the questionnaire was 0.76 and for Section 2 the Chronbach alpha coefficient was 0.66. Good internal consistency is approximately 0.80 and thus, the questionnaire as a whole and section 1 of the questionnaire show a reasonable level of internal consistency but for section 2, the internal consistency is moderate. This is probably due to the nature of the construct being investigated: positive and negative emotions and emotional regulation processes. According to Nunally (1978) a Cronbach alpha coefficient of 0.7 is an acceptable level.

3.9.3 Inter-Coder Reliability for Coding the Qualitative Data

In order to ensure reliability of the categorization and coding process, checks were made for intercoder reliability. Following Allan's (1996) approach, the following formula was used to calculate intercoder reliability:

$$\% \text{ reliability} = \frac{\text{(number of agreements)}}{\text{(number of agreements + disagreements)}} \times 100$$

(Miles & Huberman, 1994:64)

The codes assigned to the interview and open-ended questionnaire data were discussed with a colleague. The colleague is a TEFL (Teaching English as a Foreign Language) professional and is very familiar with both the context and the type of test-taker. The colleague then coded the data and the resulting intercoder reliability was 100%. There was total agreement between the two coders because the categories were very clearly defined and the data was explicit and did not yield multiple interpretations. The codes used to analyze the think aloud data were discussed with the same colleague and the colleague then coded the think aloud data.

The researcher then analyzed the reliability of responses for each of the twelve think aloud transcripts and the resulting average intercoder reliability was 87%. The responses of the two coders were compared in terms of the variety of codes used and not in the frequency of occurrence of strategies. This is a reasonable level of reliability as suggested by Miles & Huberman (1994) who recommend this level to be around 90%. Given the nature of the data being coded, a higher inter-coder reliability would be difficult to attain. The intercoder reliability ranged from 77% for the transcript of Test-taker L to 93% for Test-taker D. The two coders then discussed the disagreements and consensus was reached on each respondent resulting in 100%. This was a time consuming process but it was important to ensure reliability of coding of the data. The main areas of disagreement were not coding for the strategies of “skipping a question that is not understood and leaving the answer blank” and “never leaving an answer blank”. Other areas of disagreement varied in both test-taking strategies and emotional regulation processes in the twelve transcripts.

3.10 ETHICS

Throughout all the stages of this study, the researcher was aware of the ethical guidelines for conducting research and has demonstrated explicit conformance to ethical research practice. The researcher paid particular attention to respondents’ and test-takers’ rights to be treated with dignity and respect and ensuring the confidentiality of any information received. This researcher has also strived to be as transparent and accurate as possible by including all relevant details and comprehensive explanations of the context, respondents, instruments and methods used and the findings obtained. The researcher has included further details in the appendices of this thesis.

This thesis conformed to the ethical guidelines produced by the British Educational Research Association (shown in Appendix 14). The researcher abided by an ethic of respect of individuals, knowledge and quality of research all throughout the study. All the data has been reported ethically and accurately. The data, findings and conclusions are reported in great detail to avoid any misrepresentation and to allow other researchers to understand and interpret this data. All throughout the study, the researcher has strived to communicate the findings in a clear and straightforward manner. No information was withheld and there was no

selective reporting of findings. The appendices include all the qualitative data collected for Research Question 1 and samples of the data obtained for Research Questions 2 & 3. The statistics on the quantitative data have been reported on in detail.

All respondents were informed in advance of the purpose of the research. All those who participated in the study had volunteered and had given prior consent. Any respondent who did not want to continue a particular task had the right to leave (this occurred while administering the TTSQ when some test-takers could not stay after the test to complete the questionnaire). All the respondents were made aware of their right to be anonymous. In case of the TTSQ where respondents had to write their names (to enable the researcher to identify the proficiency level of each respondent) anonymity was not achieved, however, participants were made aware of this as indicated in the instructions given when collecting the data. Because the researcher is an Egyptian, she is aware of the cultural, religious and gender differences within the research population when conducting the study and reporting on the results.

3.11 EVALUATING THE RESEARCH DESIGN

The validity and reliability of various aspects of the methods used in this study have been discussed in detail in sections 3.8 and 3.9 of this chapter. In this section, the research design will be evaluated as a whole. In this mixed methods study both quantitative and qualitative approaches have been given equal weight. Both approaches have been really integrated in a multi-methods design (Flick, 2002). For example, the Language Testing Processing Model was empirically investigated using the think aloud data and the think aloud data was used to design the Likert-type questionnaire. The insights gained from the think aloud data helped the researcher interpret label the factors in the factory analysis conducted on the quantitative questionnaire data.

Miles & Huberman (1994:278-79) proposed an interesting framework for evaluating a mixed methods study (Appendix 12 shows the criteria and the relevant queries for each). The following Table (9) shows how the research design of this study matches the criteria proposed.

Table (9): Evaluation of research design of the thesis

Criteria	Evidence
Objectivity / Confirmability	<p>The study's methods and procedures are described in detail in this chapter illustrating how the data were collected, analyzed and presented. All interpretations of the data are clearly linked with exhibits of displayed quantitative data or direct quotes from qualitative data. All the qualitative data for Research Question 1 and samples of data collected for Research Questions 2 & 3 are included in the appendices. All the quantitative and qualitative data are available for reanalysis by any researcher.</p> <p>The researcher is aware of personal assumptions, values and biases. She considers the fact that she is a native speaker of Arabic and is very familiar with the context to be an asset in this study. This was of particular significance when conducting the interviews (she was able to elicit useful data) and interpreting the think aloud data (she was able to interpret the tone of the respondents and not just the words on the tapescripts).</p>
Reliability / Dependability / Auditability	<p>The research questions are clear and it has been shown that the research design is congruent with them. The research paradigm has been discussed in great detail in section 3.3.</p> <p>The findings show meaningful comparison across data sources as seen when comparing the results of the think aloud and TTSQ data.</p> <p>The data collection methods were congruent with the research questions.</p> <p>The coding checks for the qualitative data are described in detail in section 3.9.</p> <p>Data quality checks were made: think aloud tapes that were not clear or did not contain sufficient introspective data were discarded; TTSQ questionnaires that were not 80% were discarded.</p> <p>Multiple observers' accounts did converge as expected when defining the construct of language proficiency by students and teachers. Some convergence occurred between the think aloud and TTSQ data.</p> <p>A colleague reviewed the data and research. She is also an Egyptian who works in the same field and context as the researcher and thus, is familiar with the issues being explored.</p>
Internal Validity / Credibility / Authenticity	<p>Detailed and context-rich descriptions have been provided to the extent possible. For example, the think aloud data for each test-taker has been described in great detail and comprehensively.</p> <p>Triangulation did produce generally converging conclusions as seen when defining the construct of language proficiency by students and teachers. Some convergence occurred between the think aloud and TTSQ data.</p>

Table (9): Evaluation of research design of the thesis (continued)

Criteria	Evidence
Internal Validity / Credibility / Authenticity	<p>The data were linked to prior or emerging theory. For example, think aloud data was analyzed in accordance with the proposed LTP model which was postulated based on the literature review. The construct of language proficiency was based on the perceptions of the teachers and students.</p> <p>Areas of uncertainty were clearly identified when interpretations were not clear (Chapters 5 & 6).</p> <p>Negative evidence was sought. When the think aloud data failed to support a component of the LTP model, the model was revised.</p>
External Validity / Transferability / Fittingness	<p>All the respondents have been fully described in section 3.5 for comparison with other studies.</p> <p>Generalizability and sampling have been discussed in detail in section 3.7.</p> <p>Findings are discussed in detail for readers to assess the potential transferability in chapters 4, 5 & 6.</p> <p>The conclusions included in the LTP model presented in Chapter 7 are sufficiently generic to be applicable in other settings.</p> <p>All or samples of the qualitative data (narrative sequences) are included in the appendices.</p> <p>Chapter 7 includes a section on recommendations for further research which suggests settings where the findings could be further tested.</p>

It should be noted that while the researcher has presented sufficient detail for readers to assess the potential of transferability or relatability of this study to other contexts, there are specific considerations to be taken into account. The study was conducted on a specific group of informants: native speakers of Arabic studying English in an EFL context. Given the nature of this study which focuses on the perceptions of proficiency and strategies and emotions in a test-taking context and the nature of Arabic as a native language, it would be difficult to generalize or transfer specific findings to contexts where the native language of the respondents is Chinese or Spanish for example. In this case, only specific aspects of the study would be transferable or relatable such as the methodology used and the LTP model.

3.12 ANALYSIS OF DATA

3.12.1 Analyzing Quantitative Data

All of the quantitative data was analyzed using the statistical package SYSTAT version 10 (2000) with data imported from an EXCEL file.

3.12.1.1 Factor Analysis

Regarding the quantitative data, factor analysis was used to reduce and analyze the responses to the Likert-type questionnaire. Factor analysis is used to identify relationships between variables and to classify them. The Principal Components Analysis (CA) factor analytic procedure was used with Varimax rotation. The number of factors extracted were based on the following two criteria: minimum eigenvalues of 1.0 (the Kaiser criterion) which means that each factor accounts for at least 3% of the total variance and each factor to contain individual items with a minimum loading of 0.45 (as recommended by Hair et al, (1998)). Another method used to extract factors is the scree test which is a graphical method that is a simple line plot of the eigenvalues. The Kaiser criterion sometimes retains too many factors while the scree test sometimes retains too few. For each factor analysis (items in sections 1 and 2 of the TTSQ) the researcher examined the results of both the Kaiser criterion and the scree test (the scree tests for the two factor analyses are included in Appendix 15) and decided on the results that were interpretable and made sense. In both factor analyses the factors were extracted based on the Kaiser criterion.

In order to maximize the reliability of the correlations between the factors it is important to have a sufficiently large sample size. Various rules of thumb have been proposed; and, according to Hair et al (1998) the generally accepted norm is that the observation to variable ratio should range from 5 to 10. In this study, the number of observations was 497 and the number of variables was 78 resulting in a ratio of 6.4 which falls within the acceptable norm indicate sufficient sample size for conducting the factor analysis.

3.12.1.2 Analysis of Variance (ANOVA)

In order to compare the means of test-taking strategies and emotional regulation processes across different levels of test performance, several one-way ANOVAs were computed. A one-way ANOVA is used to test the effect of a single

categorical independent variable (proficiency level) on a single interval dependent variable (test-taking strategies and emotional regulation processes identified from the factor analysis). ANOVA focuses on F-tests of significance of differences in group means. When computing many comparisons involving the same means, the probability that one or all of these comparisons will be statistically significant increases. This is referred to as experimentwise error (a type I error for a set of statistical tests in one study). Therefore, in order to avoid an experimentwise error, a post hoc test was used to determine where the four groups differed. As stated by Wilkinson (1990:488) a post hoc test “protects you from declaring pairs of means different when they could differ by chance.” The Scheffe test which is the most rigorous post hoc test was used to compare the means of the four groups. Because 23 different ANOVAs were carried out, the level of significance was set at $p < 0.002$ for each measure in order to preserve an overall level of 0.05 for the analysis as a whole.

3.12.1.3 Correlations

Pearson product moment correlation coefficients were computed between test-taking strategies and emotional regulation processes / strategies. An acceptable significance level was deemed to be $p < 0.05$ for a two-tailed test.

3.12.1.4 Discriminant Analysis

The value of discriminant analysis (DA) is the way it can be used as an exploratory tool. DA is used to study group differences on multiple variables. It is used to identify which variables discriminate between two groups. DA first determines whether groups differ with regard to a variable and then uses that variable to predict group membership. Forward stepwise analysis is used where a model of discrimination is built step-by-step. At each step all variables are reviewed and evaluated to determine which one will contribute most to the discrimination between groups. That variable is then included in the model, and the process is repeated. Multiple DA is an extension of DA which is used to classify a categorical variable which has more than two categories (four proficiency levels) using as predictors a number of independent variables (15 test-taking strategies and 8 emotions and emotional regulation processes). The purpose of MDA is also to investigate

differences among groups, to determine the most parsimonious way to discriminate among groups discard variables that do not contribute to group differences, classify cases into groups and find out whether they are classified correctly or not (Pedhazur, 1973).

A measure of difference between the groups on the variables is the Wilks' Lambda. Wilks' Lambda varies from 0 to 1 and the smaller the lambda, the greater the differences. Thus a zero means that group means differ and 1 means that all the group means are the same. Standardized canonical discriminant function coefficients are used to identify the importance of the independent variables and as a rule of thumb it is suggested that coefficients equal to and greater than 0.30 are significant. A classification matrix is a table that is used to assess the performance of DA. The rows are the observed categories of the dependent variables and the columns are the predicted categories (Pedhazur, 1973).

A 2-group DA was conducted for high / low anxiety and high / low confidence groups of test-takers with the dependent variables: 15 test-taking strategies and 8 emotions and emotional regulation processes. A multiple DA was carried out for the independent variable: 4 levels of proficiency with the same dependent variables.

3.12.2 Analyzing Qualitative Data

Both inductive and deductive methods were used to analyze the qualitative data obtained. Inductive methods were used to explore the data related to the definition of the construct of language proficiency in the specific Egyptian context. There were no a priori definitions and the research was seeking to identify patterns. Inductive methods were also used to explore emotions and emotional regulation processes involved in test-taking of Egyptian EFL learners where the concepts and constructs have not been clearly delineated. Inductive methods were used to explore the relationship between test-taking strategies and emotional regulation processes used in the test-taking process. This process was investigated inductively and the think aloud data was analyzed with no a priori assumptions regarding the relationship between cognitive strategies and affective factors. The results from this initial data analysis were then analyzed deductively and used to further refine the proposed Language Testing Processing (LTP) model. Deductive methods were also used to

categorize the test-taking strategies of Egyptian EFL learners. Different test-taking strategies identified by previous research in addition to the strategies identified from the think aloud data provided the basis for constructing a taxonomy of test-taking strategies. Thus, part of the taxonomy was based on previous research and was used to analyze the test-taking strategies that emerged from the think-aloud protocols.

When analyzing the qualitative data obtained from the semi-structured interviews, open-ended questionnaires and the think aloud protocols, the researcher categorized or coded the data using as many categories as possible. The aim was to identify and describe the patterns found as reported by the respondents and then seek to understand and explain these patterns. The issue of coding is very important. Like analytic induction, it is an iterative process. The “researcher engages in a constant dialogue with the data until an end state of ‘adequate description’ is achieved (Byrne, 2002:148).” Basically, the process involves developing themes from the data and then examining through the data once again to find out whether these themes are representative or applicable. The whole process is inductive because the themes emerge from the data or from empirical investigation. In order to ensure reliability of the categorization and coding process, checks were made for inter-coder reliability as discussed above.

3.12.2.1 Analyzing Research Question 1 Qualitative Data

A record of the responses to the Student Interview Protocol questions for each interview and the responses to the open-ended questionnaire are shown in Appendices 7 & 8. The raw data which consists of the interview responses and the responses to the open-ended questionnaire were partially processed and transcribed onto disk. The data was then organized by grouping all the responses to each interview question or questionnaire item. The data for each question was then printed out on paper for analysis. The first step was to read and re-read the data in order to develop a detailed knowledge of all the responses and to facilitate seeing patterns. The researcher then looked for repetitions and noted these down since this is the basis for devising codes which is a key process in data analysis. “The essence of qualitative data analysis of any type is the development of a typology of categories or themes that summarize a mass of narrative data” (Tashakkori & Teddlie, 1998:119). The researcher then re-read the data several times and finalized the categories and assigned codes. These categories emerged from the data and were not

determined a priori or obtained from previous research. The different categories on the paper transcripts were color coded. At this stage also, the researcher noted unusual quotes that will be later included in the narrative on the qualitative data analysis. Tables (10) and (11) show the codes that were assigned:

Table (10): Codes for different contexts of language use

Code	Interpretation
W	use of English in the workplace
ST	use of English for studying
T	use of English for travel
S	use of English for social purposes with foreigners

The researcher read through the data, identified the different contexts that were mentioned where English was used and assigned codes. The researcher then read through the data to identify the different components of English language proficiency mentioned and assigned codes accordingly.

Table (11): Codes for components of language proficiency

Code	Interpretation	Code	Interpretation
F	Fluency	L	Listening
C	Communication	R	Reading
P	Pronunciation	W	Writing
O	Oral expression	Sk	Four skills: listening, speaking, reading and writing
A	Accuracy, grammar, sentence structure	V	Vocabulary, words, meaning of idioms, phrases

Reflecting the pragmatic approach of this study, the emerging themes from the qualitative data were analyzed by quantifying the data and reporting on the occurrence of the codes by frequency counts. In order to examine the relationship between contexts and components of proficiency, the group of codes assigned to each informant was studied and the interrelationships identified.

3.12.2.2 Analyzing Research Question 2 Qualitative Data

All of the taped think alouds or verbal protocol data were directly translated from colloquial Arabic and documented in English. The raw data consists of a

record of the twelve tapescripts and four samples are shown in Appendix 10. The data was analyzed in two stages. In the first stage an inductive approach to analyzing the data was adopted and each tapescript was read several times in order to obtain an overall view of the cognitive and metacognitive strategies used and the emotions expressed by each test-taker. The researcher then documented an overall description of the process of each test-taker's test performance in each section of the test. A deductive approach was then applied and the data obtained from each test-taker was further analyzed in light of the researcher's proposed model of self-regulation during test-taking (the LTP model). The outcome of the first stage of the verbal protocol analysis is the identification of test-taking strategies and emotional regulation processes as generated from the think aloud data.

In order to provide an overall description of the test-taking processes used by each test-taker for each section of the test the researcher two main sources of data were used: the tapescripts of each of the twelve think alouds and the scored answer sheets (both the multiple-choice and composition) of each test-taker. After reading and rereading each tapescript, the researcher analyzed the responses and then related these interpretations to actual test / item performance (test-takers' scores). Four samples of test-takers' actual responses are shown in Appendix 11. In order to test the proposed language testing processing (LTP) model (page) and examine the interplay between test-taking strategies and emotional regulation processes used by Egyptian adult EFL learners while taking a test, detailed descriptions of each of the twelve (12) test-takers' responses were provided. In order to find out whether or not the LTP model does apply, the descriptions of each test-taker are presented in terms of how they assessed their capability in performing the task (a goal orientation must be present for emotions to emerge), their affective responses, the strategies they used to handle the task and to deal with their emotions and their subsequent behaviors. "Thus, strategies are inferred from verbal protocols by examining the protocols for particular sequences or patterns of processes" (Green, 1998:71). Test-takers' actual responses were examined in order to differentiate between contributory and non-contributory strategies. Contributory strategies are those that lead to the correct choice of answer while non-contributory strategies are those that do not lead to the correct response. As discussed in section 3.6.3, in order to ensure that the test was relevant and important to the test-takers, they were provided with certificates of their test results.

Given that a lot of work has been done in the area of test-taking strategies and could not be disregarded and the nature of qualitative data which is based on the perceptions of particular individuals participating in the study at a particular time, it was important to go back to the literature to identify possible strategies or emotions that did not occur or emerge from the data. In the second stage of data analysis, the strategies identified from the think aloud data were compared to the test-taking strategies previously identified from the literature and an overall taxonomy was designed. The researcher then used the taxonomy to code the think aloud data on item responses for each test-taker. This step was important to validate the strategies and emotions identified from the data and to identify further patterns in the data. Table (12) shows the list of test-taking strategies obtained from the think aloud data only.

Table (12): Test-taking strategies obtained from the think aloud data

Strategies for Taking a Multiple Choice Proficiency Test Not Specific to One Skill

- 1. Inserting each option one at a time in the question.
- 2. Translating the question and the options.

Specific Strategies for Writing

- 3. Reads the prompt in English only.
- 4. Reads the prompt first in Arabic and then in English.
- 5. Reads the prompt first in English and then in Arabic.
- 6. Reads the prompt in Arabic only.
- 7. Rereads the prompt.
- 8. Formulating ideas in Arabic and then translating them into English to write them down
- 9. Using words from the prompt.
- 10. Formulating ideas in English.
- 11. Rewriting words or phrases.

Specific Strategies for Listening

- 12. Reading the options of the next question before listening to the question on tape.
- 13. Using knowledge of grammar to answer the question.

Specific Strategies for Reading

- 14. Rereading the whole passage.
 - 15. Rereading parts of the passage.
 - 16. Reading the passage and then summarizing the ideas in Arabic.
 - 17. Pronouncing or sounding out words to find their meaning.
-

Table (13) shows the list of codes used to further analyze the think aloud data. This list of codes was generated based on two sources. The first source was

the list of test-taking strategies that were obtained from the think aloud data and did not occur in the literature (as shown in Table 12) and test-taking strategies obtained from the literature. All these strategies were combined into one list and each test-taking strategy was assigned a code to be used in analyzing the think aloud data a second time.

Table (13): List of codes to analyze the think aloud data: Test-taking strategies

CODE	TEST-TAKING STRATEGY
TTS1	Reading the questions and options before choosing one
TTS2	Stopped reading options when they got to the one that seemed correct.
TTS3	Selecting an option by eliminating the other 3 options
TTS4	Translating the question and the options.
TTS5	Inserting each option one at a time in the question.
TTS6	Making an educated guess using background knowledge
TTS7	Trying to produce their own answer to the question before looking at the options provided
TTS8	Rereading the questions and options for clarification
TTS9	Postponing dealing with a question or selecting a given option until later.
TTS10	Skipping a question that is not understood and leaving the answer blank
TTS11	Guessing without any particular considerations
TTS12	Changing responses when appropriate
TTS13	Selecting an option that is longer / shorter than the others
TTS14	Looking for an option that seems to be different from the others
TTS15	Running out of time without trying all the questions
TTS16	Trying to finish the test as fast as possible
TTS17	Previewing or surveying the whole test
TTS18	Monitoring time
TTS19	Watching to see when other students finish the test.
TTS20	Reading the instructions carefully.
TTS21	Never leaving an answer blank
TTS22	Going back and review or check answers.
TTS23	Getting stuck on one question for a long time.
	Specific Strategies for Writing
TTS24	Reads the prompt in English only.
TTS25	Reads the prompt first in Arabic and then in English.
TTS26	Reads the prompt first in English and then in Arabic.
TTS27	Reads the prompt in Arabic only.
TTS28	Rereads the prompt.
TTS29	Formulating ideas in Arabic and then translating them into English to write them down
TTS30	Using words from the prompt.
TTS31	Formulating ideas in English.
TTS32	Rewriting words or phrases.

Table (13): List of codes to analyze the think aloud data: Test-taking strategies (continued)

CODE	TEST-TAKING STRATEGY
	Specific Strategies for Listening
TTS33	Reading the options of the next question before listening to the question on tape.
TTS34	Using knowledge of grammar to answer the question.
	Specific Strategies for Reading
TTS35	Reading the passage first
TTS36	Reading the questions first before reading the passage.
TTS37	Rereading the whole passage.
TTS38	Rereading parts of the passage.
TTS39	Translating relevant parts of the passage for understanding.
TTS40	Reading the passage and then summarizing the ideas in Arabic.
TTS41	Guessing meaning of an unknown word from the context
TTS42	Skipping unknown words.
TTS43	Pronouncing or sounding out words to find their meaning.
TTS44	Using knowledge of grammar to answer the question.
TTS45	Looking for portion of the passage that the question refers to and then look there for clues to the answer
TTS46	Matching material from the passage with material in the question and in the options
TTS47	Selecting an option because it appears to have a word or phrase from the passage in it
TTS48	Selecting an option based on understanding the passage read
TTS49	Getting clues from answering one question that were helpful in answering another question

In order to analyze the think aloud data a second time all the emotions identified from the literature and from the first stage of the think aloud data analysis were combined and a list of emotions, assessment (cognitive-appraising processes), task-focusing and emotion-focusing processes was created. The behavioral effect of negative emotions experienced during the test-taking process such as giving up was also included. Since the outcomes of Research Question 2 leads to the design of the Likert-type questionnaire used to investigate Research Question 3, affective strategy #9 was added to identify the goal orientation of test-takers i.e. whether taking the test was relevant or not. All of these emotions, emotional regulation processes and behaviors were combined into one list and each item was assigned a code to be used in analyzing the think aloud data a second time.

Table (14) shows the list of codes used to further analyze the think aloud data. The emotions identified from the think aloud data are shown mapped onto the table between brackets.

**Table (14): List of codes to analyze the think aloud data:
Emotional regulation processes**

CODE*	EMOTIONS & EMOTIONAL REGULATION PROCESSES
AFF1	Tried hard on the test.
AFF2	Did well. (confidence)
AFF3	Test or item was easy.(confidence)
AFF4	Test / item was difficult.(complaining / anger)
AFF5	Test / item was confusing. (confusion)
AFF6	Felt like cheating.(wishful thinking)
AFF7	Felt prepared. (confidence)
AFF8	Knows what to do.(confidence)
AFF9	Felt it was important to do their best.
AFF10	Had enough time to finish the test.
AFF11	Enjoyed taking the test. (pleasure)
AFF12	Felt that taking the test was a challenge.
AFF14	Felt relieved when the test was over.(relief)
AFF15	Felt nervous during the test. (anxiety)
AFF16	Had difficulty in concentrating.(complaining)
AFF18	Dreads taking tests because he/she knows more than the test will show.(high anxiety)
AFF19	Felt tired during the test. (tiredness)
AFF20	Gave up because the test was too difficult.(helplessness)
AFF21	Got tired and started answering without reading the question.(helplessness & tiredness)
AFF22	Because of nervousness forgot the things that they usually know.(disappointment)
AFF23	Was not sure of the correct answer.(confusion)
AFF24	Felt the test was interesting.(pleasure)
AFF25	Felt bored while taking the test.(boredom)
AFF26	Taking the test was a pleasant experience.(pleasure)
AFF27	Felt frustrated because there was not enough time.(frustration)
AFF28	Felt it was a good idea to have the instructions in both Arabic and English. (approval)
AFF29	Felt surprised at having the writing prompt in both Arabic and English.(surprise)
AFF30	Disapproved about having the writing prompt in both Arabic and English. (disapproval)
AFF31	Felt it was a good idea to have the writing prompt in both Arabic and English.(approval)

* The numbering of the codes is not correct (after having coded the data, 13 & 17 were discovered to be missing).

The TTSQ qualitative data were in Arabic. Initially, the data were typed in Arabic and then translated into English (see Appendix 13). Similar to approaches described above, the researcher read and reread the data several times and identified the underlying patterns and codes were determined. The data were then coded and reported on.

3.13 SUMMARY

This chapter has provided the rationale for the design of the study. It is located within the pragmatist paradigm and the research questions or issues in this study are investigated using a combination of qualitative and quantitative methods. The data collection instrument design, piloting of instruments, data collection procedures and analysis of qualitative and quantitative data are described in detail. The respondents who participated in the study are also described. The approach adopted in this design seeks to yield data that are both valid and reliable as demonstrated in this chapter. Several features are included in the design of the study to ensure internal and external validity (triangulation is one of these design features) as well as reliability (one example is intercoder reliability). The research design ensures that the data obtained is sufficiently meaningful to represent the constructs being described and is sufficient to provide verification of findings that emerge. The quantitative data is analyzed using several statistical tools and both inductive and deductive methods are used to analyze the qualitative data obtained. Considerations of quality control in all stages of conducting the research are also demonstrated.

CHAPTER 4:

RESULTS & FINDINGS FOR RESEARCH QUESTION 1

4.1 RESEARCH QUESTION 1

How is the construct of foreign language proficiency defined in a specific Egyptian EFL context?

- 1.1 What are the different contexts of language use of Egyptian adult EFL learners?
- 1.2 How do adult EFL learners in a specific context in Egypt define language proficiency?
- 1.3 To what extent does the construct of language proficiency vary according to the context of language use?

In order to answer Research Question 1, the construct of language proficiency was investigated from the both the learners' (N=36) and the teachers' (N=41) perspectives in the Center for Adult & Continuing Education at the American University in Cairo.

4.2 FINDINGS ON MOTIVES FOR STUDYING ENGLISH AND DIFFERENT CONTEXTS OF LANGUAGE USE

The reasons for studying English were varied, however, the most frequently occurring responses that were identified and coded were related to the use of English: in the workplace, for studying, for travel and for social purposes. The first reason mentioned by both groups of respondents for studying English was related to the workplace, either to find a better job / better career opportunity or to improve their current job. Several respondents had jobs that required the use of English. An air transport controller stated that "I need English to deal with requests from foreign companies and embassies and to negotiate with delegations." A student who had a job in export and import stated that "I deal with customers over the telephone and all the numbers, documents and certificates I work with are in English." Another respondent works as a hostess on a coach or bus and noted that "I sometimes use English with

foreigners or tourists going to Hurghada, Luxor or Aswan.” One respondent stated that computers will be introduced at his place of work and he would like to be prepared. Several students specifically mentioned wanting to work for foreign organizations within Egypt or abroad. A graduate of the Institute of Koranic Reading stated that he needed to learn English in order to explain the meaning of the Koran to people in other countries. Another respondent stated that “my fiancé works in Oman where I will need English to find a job.” A first year student in the faculty of commerce stated that “English is required for any job now. English is a must.” The majority of the teachers (36/41) stated that the main reason that Egyptian adults study English is to find good jobs or get better jobs. One teacher stated that “university students come to complement their academic studies once again in the hope of finding a good job.” Another teacher noted that “I think they study English because it’s a must that they use it to get a decent job.” Several teachers mentioned that because of unemployment, knowledge of English may give these adults an edge over others applying for certain jobs. A teacher noted “I think Egyptian adults study English because the rate of unemployment is increasing, thus some knowledge of English may given privilege above other applicants for certain jobs ...” Some teachers stated that studying English will help these students to improve their career and may get them promoted to better positions. A teacher stated “their main objective is a certificate... helps them in getting better jobs if they are at entry level. Once in a company English does not matter to them unless they are competing with another person or move into export department or to further secure their jobs.” Thus, knowing English will enable adults to compete in the job market both at entry level and within their current work situations.

The second most important reason was for academic purposes meaning that the EFL learners needed English for their studies. This was mentioned by thirteen (13) respondents. Two students stated that studying English would help them with their current undergraduate studies in the faculty of arts, English section and one student stated that “I find studying English literature to be difficult because I need to improve my English.” Three students, enrolled in the faculty of commerce who did not need English for their studies, were taking courses in order to be admitted into the non-credit accounting or computer diplomas at CACE. Several postgraduate students were

studying English in order to complete their MA or PhD degrees Egypt or to obtain a particular TOEFL score in order to complete their postgraduate studies abroad. A respondent who is doing postgraduate studies in sociology noted that “English is essential for higher studies in reading references and professors look at the English references in the bibliography.” Twenty-one (21) teachers stated that English is needed by students to continue postgraduate studies in Egypt or abroad. Most of them mentioned that English is required for computer studies or to learn and work on computers. Other fields of study mentioned were business and preparing for the TOEFL examination. Several mentioned that university students study English to complement their studies in the national universities, and one teacher specifically noted that students in the faculties of English literature or education needed to improve their English. One teacher stated that English has “become a sort of gateway that they want to have wide open on the different fields of knowledge”. The language skills identified for study purposes were writing a thesis in English, reading references, translating articles, and attending lectures in English.

Another motive for studying English was a travel motive (mentioned by 7 students and 13 teachers), either to visit relatives in the US or the UK or to emigrate to the US or Canada. A student noted that she was studying English because “I hope to travel to the US to visit my sister there.” A teacher stated “they want to travel and know that good English would facilitate their life abroad.” Three (3) students mentioned studying English in order to help their children and twelve (12) teachers stated that adults study English to help children with their studies at school, especially if these children are enrolled in language schools. A teacher stated that “some young mothers study English to help their kids who go to language schools.”

Another reason that Egyptian adults study English according to 14 of the teachers is a social motive. They believe that students study English to improve their cultural status or enhance their social standard. One teacher stated that learning English gives students ‘a sense of importance and individuality in the community.’ Some teachers believe that students attend English classes to socialize, to meet friends and colleagues or to find a partner (husband or wife). One teacher stated “they come dressed to kill in the hope of gaining friends or even a mate” and another stated that “some

young ladies and divorced women think it's a good chance to make social relations." Some teachers noted that for some students coming to class is a useful way to pass time or to get away from work for a few hours.

Several students (five) had no immediate reason for studying English but perceived that it was important for them to learn or improve their English for the future. Some of them were worried about forgetting their English and wanted to maintain their current level. One subject stated that "speaking English in class is the only opportunity I have for practicing the language".

The importance of English was emphasized by all the interviewees and two subjects went as far as stating that "someone who does not know English is ignorant and uneducated and not keeping up with the times". They added that "knowing English is not a luxury but a basic need in order to keep up with the West". This underscores the status of English in the society.

Table (15) shows the frequency counts for the different contexts identified by both groups of respondents:

Table (15): Contexts of language use identified by students and teachers

Context	Frequency Count	
	Students	Teachers
Use in the workplace	23 (64%)	38 (93%)
For studying / academic purposes	16 (44%)	21 (51%)
For overseas travel	7 (19%)	13 (32%)
For social purposes	7 (19%)	14 (34%)

The principle context of language use identified by the two groups of respondents was career / work-related. Within the career context the pattern that emerged from the student data was a continuum which ranged from using English on a daily basis to using it occasionally. English was used mainly for dealing with foreigners and / or computers. Almost all the teachers (with the exception of three) identified the work context to be the main situation where students would use English. Some noted specific organizations such as foreign companies, banks, hotels, restaurants and the

airport. Some teachers also listed specific professions that require the use of English such as tourism, sales, journalism, public relations, air traffic control, doctors, engineers, secretaries and those who work on computers.

The second frequently occurring context of language use was the academic context. The academic context can be defined as enhancing the ability to study or conduct research where the medium of instruction is English. It is important to note here that the medium of instruction is English in several Egyptian universities which reflects the significance of academic English. Sixteen (16) students and twenty-one (21) teachers mentioned the use of English by EFL learners for their current studies or pursuing higher studies.

The third context that emerged was the use of English for social purposes. This includes the use of English for traveling overseas (mentioned by seven students and thirteen teachers), to communicate with foreigners, to watch English video films or movies and to use some words with friends.

Fourteen (14) of the thirty-six (36) respondents stated that they currently do not use English but are studying English for some future reason. This unknown future context can be defined as career oriented or academically oriented or personal enhancement. This is the context where subjects do not actually need or use English in their current situations but may do so in the future. Four (4) teachers believed that outside the classroom, opportunities for Egyptian adults to use English are extremely limited and as one teacher stated "other than the classroom chances are nearly non-existent" to encounter English speaking foreigners.

Only seven (7) subjects had very specific language targets they were aiming for such as the university professor who was organizing a local conference with English as the primary language. She specifically needed to improve her speaking skill, focusing on grammar and vocabulary, within a three-month period. Another wanted to achieve a specific score on the TOEFL within six months and a third wanted to improve her writing within six months to complete her MA degree. A student was studying English in order to sit for the Certified Public Accounting (CPA) exam. Three subjects were aiming to reach a particular level (level 12) at CACE in order to be admitted into a non-credit certificate program.

Twelve (12) subjects noted that they would study until the end of all the courses offered and eleven (11) did not give a timeframe. Three stated that they would continue if they found the course beneficial. They felt that as long as they were progressing, the classroom was the best learning environment. Eleven (11) subjects stated a timeframe ranging from six months to two years and one subject mentioned a period of three years. Therefore, only one third of the subjects specified how long they intended to study English.

To conclude, both the teachers and learners were in agreement regarding the two main motives for studying English which were work-related and for academic purposes. The teachers identified a third reason which was a social motive related to the cultural status of English. This was not directly stated by the learners, however, it could be inferred from their responses to other questions. At least one third of the learners did not need or use English in their current situations and did not seem to have specific linguistic targets or timeframes for studying English. Furthermore, they all emphasized the importance of knowing English or maintaining their current level in order to be successful in life. Teachers tended to overemphasize the travel motive, whereas only seven learners mentioned it as a reason for studying English. Similarly, teachers strongly emphasized the motive to help children with their school studies.

Regarding the contexts for use of English, both teachers and learners agreed that the two main contexts were the work and academic contexts. The teachers mentioned a social context for use of English which was referred to only by a few of the learners. About one third of the learners did not need or use English in their current contexts and were studying English for personal development or for a future job or academic context.

4.3 FINDINGS ON DEFINING THE CONSTRUCT OF LANGUAGE PROFICIENCY

In order to define language proficiency students were asked to describe the difficulties they face in using English and both students and teachers were asked to describe the proficiency of a person at the top of the English proficiency continuum. Regarding the difficulties students faced in each of the four language skills, almost all the respondents (34 of the 36) felt that they could not express themselves orally and that

they had problems with listening comprehension. Examples of difficulties mentioned were that they could not find words easily, could not cope with the fast pace of a conversation or a tape, or understand the different accents. Some of the students' comments were "I find it difficult to find the right word", "I can't follow a conversation", "I can't speak with a foreigner", "I can't say what I want in English", "I find listening to be the most difficult. I can't transform English words to Arabic fast and I find it difficult to listen to foreign songs" and "Listening is my weakest skill especially when listening to Voice of America or BBC". Five respondents specifically mentioned that they found it difficult to understand the American accent whereas they perceived the British accent easier to understand. One student stated "I find listening to an American accent difficult" while another noted "British is slower and easier". It is interesting to note here that the background of the educational system in Egypt is British. Four (4) students mentioned pronunciation as a difficulty in speaking.

Ten respondents indicated that they had difficulty with writing. The main problems identified were in expressing their ideas and spelling. One student mentioned that she had a problem in writing in both Arabic and English. Only five (5) respondents mentioned that they had difficulty in reading. One student stated that when she took the exam she got very depressed because she knew nothing and she felt embarrassed to come and get her placement test results. It is interesting to note that seventeen (17) students stated that they did not have any problem in reading and writing.

Table (16) shows the analysis of the descriptions of students and teachers of the language proficiency of a person at the top of the continuum. The responses were coded and the frequency counts are shown in the table.

Almost all students (except for six respondents who did not explicitly mention listening or speaking) and about half of the teachers included in their definition of language proficiency fluency in speaking or some aspect of oral fluency. It is seen as the ability to communicate orally, pronouncing the words clearly using a good accent which is either American or British. Several mentioned that the way of speaking is clearly understood and the speaker does not have to repeat or clarify what was said.

Table (16): Components of language proficiency identified by students and teachers

Language Proficiency Components	Frequency Count	
	Students	Teachers
Fluency	16 (44%)	23 (56%)
Oral expression	30 (83%)	21 (51%)
Pronunciation	14 (39%)	9 (22%)
Vocabulary	8 (22%)	16 (39%)
Communication	8 (22%)	15 (37%)
Accuracy	6 (17%)	19 (46%)
The four skills combined: both oral and written	6 (17%)	15 (37%)
Reading	7 (19%)	6 (15%)
Writing	8 (22%)	11 (27%)
Listening	--	2 (5%)

Fluency is the ability to speak easily without searching for words, without stopping to translate, expressing ideas quickly without thinking or hesitation as they would speak Arabic. The focus in fluency is on speed and automaticity and the person does not take long in expressing him/herself or understanding. These findings are in agreement with Meara’s (2002:404) statement that “the ability to recognize and retrieve words effortlessly seems to be a basic feature of the performance of L1 speakers and a feature that is conspicuously lacking from the performance of most L2 speakers.” Respondents also view proficiency as the ability to watch foreign films or serials without translation. One respondent stated that “being proficient means that he can participate in any conversation and it is not someone who studied English the way we did aiming only to pass from year to year”. Another student stated that to her those who are proficient in English are “graduates of language schools and simultaneous translators. Their life is all English from KG and they speak English to each other; they don’t speak Arabic.” A teacher described oral fluency as the ability to “sustain a steady flow of information with very minor pronunciation pitfalls that do not affect the flow.” Another teacher stated that the speaker is “no longer conscious of grammar or vocabulary but expresses himself and understands automatically”.

Another component of proficiency mentioned by 22% of the learners and 39% of the teachers is having a rich repertoire in vocabulary which leads to fluency and ease in

speaking. One respondent stated that “those who appear on the English TV broadcasts” are examples of people proficient in English. Another student described a proficient person as someone who has all the vocabulary available and who can understand multiple meanings of an expression or sentences with different meanings. Teachers stated that proficiency means the idiomatic use of English, the use of the English lexicon appropriately and correctly, good knowledge of vocabulary and the understanding of the meaning of idioms from context.

Both teachers (37%) and students (22%) included communication as a component of proficiency. Students described communication as dealing with others in a foreign country, dealing with ordinary people in an English-speaking community, expressing ideas simply and conveying meaning and dealing with any situation. Teachers defined communication as understanding spoken and written English in various situations, using the language in different situations, negotiating meeting, conveying self-expression and communicating freely.

The main difference between students and teachers in their definition of proficiency is the component related to structure of the language and accuracy. Almost half of the teachers (46%) mentioned the use of correct grammar, accuracy of structure, language free from grammar errors and minimal mistakes. Only 17% of the students mentioned the use of simple and correct English and the grammatical structure of sentences. Teachers because of their profession not only focus on fluency but also on accuracy in terms of correctness of language use and minimizing errors whereas students focus on using the language.

Six students defined proficiency as being excellent in the four skills and seven (19%) others focused on reading and eight (22%) mentioned writing skills. A larger proportion of teachers (37%) compared to students described proficiency in terms of the four skills and 15% mentioned reading while 27% included writing. It is interesting to note that only one teacher mentioned awareness of cultural backgrounds and none of the learners discussed any aspect related to culture.

It is clear that learners in the context of the current research perceive themselves as language users in terms of their ability to speak the language and understand those who speak. The focus is on oral communication. Teachers’ descriptions of persons at

the top of the English proficiency continuum were mainly based on fluency in the four skills (listening, speaking, reading, writing) and / or ability to express themselves with accuracy using correct grammar and appropriate vocabulary including understanding and using a wide range of idioms. All those who mentioned the skills, included speaking followed by writing. Some descriptions also included ability to communicate in different situations and with native speakers. Only nine teachers (22%) mentioned pronunciation in their descriptions compared to 39% of the students which reflects the communicative approach or perspective held by teachers, responding to the questionnaire, towards teaching English as a foreign language.

4.4 FINDINGS ON THE RELATIONSHIP BETWEEN THE CONSTRUCT OF LANGUAGE PROFICIENCY AND THE CONTEXTS OF LANGUAGE USE

In order to find out the extent to which the construct of language proficiency varies according to the context of language use, the responses to Research Question 1.1 and 1.2 were compared. On examining students' responses, it can be seen that 64% of the students stated that they needed English for the workplace and 44% mentioned that they used English for academic / studying purposes. The majority of these students (83%) defined proficiency in terms of oral fluency regardless of the different contexts they claimed they used English for. Students who stated that they need English for their current studies such as studying in the faculty of arts, English department or the institute for hotels and tourism, or studying for a Master's degree in sociology or mass communication still described proficiency in English in terms of speaking or oral fluency. If students had linked their definition of proficiency to the context of language use, it would have been expected that they view language proficiency in terms of the ability to read and write and not in terms of the ability to speak only. On examining the responses of teachers it can be seen that there is a link between the context of language use and the definition of language proficiency. The majority of teachers (93%) stated that the main context of language use is in the workplace and they defined proficiency in terms of fluency and accuracy in the four skills.

As discussed in the literature review and in light of Bachman & Palmer's (1996) model, when constructing a language test, a contextualized test design approach is far more useful than adopting different components from existing theoretical models of language proficiency. Since the aim of Research Question 1 is to empirically derive the components of language proficiency of adult EFL learners in a specific context in Egypt in order to use it to design a placement test, the following section describes the construct that was developed.

4.5 DEFINITION OF THE CONSTRUCT OF ENGLISH LANGUAGE PROFICIENCY OF ADULT EFL LEARNERS AT CACE, AUC

Based on the findings, the three main contexts of language use identified were the workplace, academic and social contexts. Regarding the components of language proficiency, although adult learners view language in terms of oral proficiency regardless of the context of language use, however, teachers present a more realistic definition of proficiency in terms of the four language skills: listening, speaking, reading and writing. Language proficiency in academic contexts cannot be defined in terms of oral fluency. Thus, the construct will be based on the contexts identified by the learners and the language proficiency components defined by teachers. Following is the definition of English language proficiency of adult EFL learners at CACE, AUC: it is the ability to communicate fluently orally and in writing in the workplace, for social purposes and to read and write for academic or educational purposes.

In the literature review various models of language proficiency (Munby, 1978; Canale & Swain, 1980; Oller, 1983; Bachman, 1990 and Bachman & Palmer, 1996) were presented and critically analyzed. The Bachman & Palmer (1996) model was found to be superior because it addressed the weaknesses found in the previous models and the approach adopted by this model is that tests should be designed to measure constructs in a given context. Without this definition of the construct, curriculum designers would not be able to design a curriculum framework or materials and test designers would not be able to design valid language placement tests. Therefore, the language proficiency construct established in this study was used as the basis for

designing a new English language curriculum and placement test in the Center for Adult & Continuing Education in the American University in Cairo.

The construct definition in this thesis is consistent with the findings of a subsequent study (Aydelott et al, 2000) conducted on a sample of 658 Egyptian adult EFL students and 74 teachers in the same program (CACE at AUC). This study provides further empirical evidence of the validity of the components of the construct being defined. The aim of the Aydelott et al (2000) study was to conduct a language needs assessment in order to design a new curriculum based on identified language needs. A questionnaire was used and both quantitative and qualitative data were collected. The conclusion was that EFL proficiency in the Egyptian context was defined as “the ability to communicate (engage in conversations) in social settings with people from different cultural backgrounds; the ability to use English (engage in both oral and written communication) in the workplace; and the ability to understand readings obtained through written or electronic media”. Therefore, this construct definition of EFL proficiency which is very similar to the definition reached based on the findings of this thesis, is the one adopted throughout this study. The English language Placement Test (EPT) was subsequently redesigned based on this view of language proficiency.

Because of the large number of students that have to be placed within a short period of time (approximately 4000 students per term) before the beginning of each term, CACE decided that the new EPT administered to all students would consist of a paper-and-pencil test. This would facilitate scoring and the results could be produced within one or two working days. The paper-and-pencil test consisted of a multiple-choice component with a listening section and a writing component. The speaking component of the test (the interview) would be administered only in cases of misplacement: students who were placed too high or too low in the English language program. Thus, for practicality and efficiency, the ability to use English to communicate orally in the workplace and in social settings would be tested for a smaller number of students when necessary.

The new CACE EPT consists of four sections: writing, listening, grammar and reading comprehension. The writing task is related to the workplace and test-takers are given a workplace situation to respond to in writing. In the listening section, the texts

are related to both social and workplace situations. In the grammar section, the items are also presented within a social or workplace context. In this section, test-takers have to complete an error identification task where they are required to select the part of a sentence that includes a grammatical error. This skill is important in the workplace context where people are required to proofread and review their own writing or that of others. In the reading comprehension section, test-takers are required to read a variety of texts related to the workplace and texts obtained from the internet.

It should be noted that the new language curriculum now consists of four levels of language proficiency (the old curriculum consisted of six stages): novice, elementary, intermediate and advanced and the EPT is designed to place students into these stages. Appendix 16 includes the descriptions of these four stages.

4.6 SUMMARY OF FINDINGS AND OUTCOMES ON RESEARCH QUESTION 1

Based on the results and analysis above, a summary of the findings related to Research Question 1 and each sub-question is presented below.

4.6.1 Research Question 1

How is the construct of foreign language proficiency defined in a specific Egyptian EFL context? It is the ability to communicate (engage in conversations) in social settings with people from different cultural backgrounds; the ability to use English (engage in both oral and written communication) in the workplace; and the ability to understand readings obtained through written or electronic media.

4.6.2 Research Question 1.1

What are the different contexts of language use of Egyptian adult EFL learners? The three different contexts of language use of adult EFL learners are: work-related, academic and social contexts.

4.6.3 Research Question 1.2

How do adult EFL learners in a specific context in Egypt define language proficiency? All subjects included in their definition of language proficiency fluency in speaking or some aspect of oral fluency. It is seen as the ability to communicate orally, pronouncing the words clearly using a good accent which is either American or British. Fluency is the ability to speak easily without searching for words, expressing ideas quickly without thinking or hesitation as they would speak Arabic, using correct sentences.

4.6.4 Research Question 1.3

To what extent does the construct of language proficiency vary according to the context of language use? The results indicated that regardless of the context of language use, adult EFL learners perceive themselves as language users in terms of their ability to speak the language and understand those who speak. The focus is on oral communication. Therefore, the construct of language proficiency does not vary across the different contexts from the point of view of the learners. Teachers on the other hand defined language proficiency in terms of accuracy and fluency in the four skills for use in the workplace primarily and for academic contexts.

CHAPTER 5:

RESULTS & FINDINGS FOR RESEARCH QUESTION 2

5.1 RESEARCH QUESTION 2

What are the test-taking strategies and emotional regulation processes used and emotions experienced by Egyptian adult EFL learners in a specific context when taking an English language placement test?

As described in Chapter 3: Research Design, in order to answer Research Question 2, twelve (12) subjects were asked to ‘think aloud’ and to verbalize their feelings as they completed a test which consists of four sections: writing, listening, grammar and reading comprehension. The test consists of ninety (90) multiple-choice questions and a writing component and requires 1.5 hours to complete.

5.2 FINDINGS FROM THE THINK ALOUD DATA: IDENTIFYING TESTING-TAKING STRATEGIES & FIT OF THE LTP MODEL

The think aloud data were analyzed and are discussed in view of the components of the proposed language testing processing (LTP) model for each section of the test: the test-taker’s appraisal of the test task demands and his / her competence to complete the task (metacognitive strategy use), the contributory and non-contributory test-taking strategies used (cognitive and metacognitive) and the positive and negative affective responses.

The metacognitive strategies used by test-takers to assess their ability to complete the test tasks are described and the test-taking strategies identified are indicated in italics within the descriptions. The affective / emotion responses for each test-taker are summarized in a table at the end of each description. At the end of the section, all the strategies identified are listed and the workings and applicability of the LTP are discussed. Further details about analyzing the think aloud data are found in section 3.10.2.2.

It is important to note that the interpretations of the think aloud data offered by the researcher are not necessarily definitive and alternative interpretations are plausible. These interpretations are influenced by the fact the researcher is a native speaker of

Arabic and moreover, very familiar with the context. While listening to the think alouds, sometimes the tone of the test-taker was indicative of a particular emotional response and accordingly influenced the researcher's description of the emotions that occurred during the test-taking process.

5.2.1 Test-taker A (Male, elementary English language proficiency stage, 18-22 age range)

5.2.1.1 Writing

In assessing his competence, the test-taker expressed doubt about his ability to write sentences. Providing the prompt in both Arabic and English gave him confidence because as he stated "it provides test-takers with words they might not know." The test-taker's main cognitive strategy used was to *express his ideas in Arabic and then use some words from the prompt to write them in English*. Although, it was clearly stated in the instructions that the formatting of the letter is not part of the scoring criteria of the writing task, however, the test-taker coped with the task by focusing on the format when writing and stated that "even if one does not know how to write sentences, one may get points for writing the date". He believed that writing a letter is easier than writing a paragraph. The test-taker wrote three sentences: one correct simple sentence and two grammatically incorrect sentences.

5.2.1.2 Listening

In Part A he was very confident, expressed impatience at the slow pace of the tape and got almost all the items (13 items) correct. He selected the correct response right away by reading the option in English or translating it into Arabic. In Part B he still expressed confidence, however got most of the items wrong (only 3 items were correct). The test-taker did not mention any negative affective responses in spite of poor performance. The strategies he used were *translating the item and options into Arabic, reading the options, reading aloud a possible answer and then changing and selecting a final answer*. At the beginning of Part C the test-taker gave up stating that "I didn't do C because it is beyond my level." Although he found the listening task to be difficult, he

approved of including it in the test and stated that it “makes the student listen a lot and study more.... It is a very good way for students.”

5.2.1.3 Grammar

The test-taker was surprised that the instructions were written in both Arabic and English. He felt confident about his grammar ability and stated that “grammar is easy in general” and predicted that the questions would be easy. However, starting the 3rd question in this section, he found the questions to be difficult. The main strategy used was to *read the item and options*. He kept repeating that the “questions are getting more difficult.” After question 60 he gave up stating that “the exam from Section C in the listening is getting more and more difficult and is higher than my ability level.” He found the instructions confusing for Part B: error identification and the main strategy used was *guessing*. He got 8 questions correct (out of 15) in Part A and none in Part B. He guessed the answer to several items which probably resulted in poor performance.

5.2.1.4 Reading

He attempted to read the first passage. He did not record any responses on his answer sheet. The self-appraisal process was probably negative. He stated that the test was becoming very difficult and he blamed the test designers stating that they should realize that the ability of the test-takers is very low i.e. lower than that of Kindergarten 1 or 2. He noted that this exam required a lot of effort and he felt disappointed. He ended by stating that the exam is very good because it will make them study more. He then attempted to read the second passage and the main strategy used was *guessing*. He also attempted to *pronounce the word ‘guarantee’* and to use the *context to guess the meaning of ‘guarantee’* but was unable to do either. He attempted the third passage and found it very difficult. He used *guessing* and was not sure of any of the answers. He ended on a frustrated note (his tone of voice) stating that “when I get to a point in my reading, I forget the previous part... Time is not enough.”

5.2.1.5 Affective Responses

The following Table (17) illustrates the emotions that emerged from the think aloud protocol of Test-taker A as well as the indicators of these emotions.

Table (17): Affective responses of Test-taker A

Emotion	Indicators from the Think Aloud Data
Approval	He stated that he was pleased with the Arabic prompt, the writing and listening tasks.
Confidence	Initially the test-taker started out very confident but gradually lost confidence towards the end of the test. He gave up after question 60.
Anxiety	He experienced anxiety several times throughout the test in all the sections: writing, listening, grammar and reading. Writing: He stated that he did not know how to write sentence. Listening: He stated that he gave up because it was beyond his level. Grammar: He repeated that questions are getting more difficult. Reading: He stated that the test is becoming difficult. He handled anxiety in several ways: persisting and guessing (in the grammar & reading sections) or focusing on the letter format (in the writing section), giving up and blaming the test designers for making the test difficult or expressing approval of the exam stating that it would make students study more (an analogy would be that the test was like taking medicine, unpleasant but good for one).
Anger	Blamed the test designers for making the test difficult
Surprise	The test-taker expressed surprise at the bilingual instructions in the grammar section.
Confusion	He found the questions to be difficult and confusing starting from Part C of the listening section.
Disappointment	He expressed disappointment in the reading section.
Frustration	He stated that he needed more time in the reading section. His tone reflected his frustration.

5.2.1.6 Fit of the LTP model

At the beginning of each test task: writing, listening grammar and reading tasks, the test-taker assessed his language ability to perform these tasks. The metacognitive strategies used were: negative assessment of language ability in a specific skill or task (the test-taker did not feel confident about his ability to write sentences), positive assessment of language ability in a specific skill or task, predicting that the task would be easy and deciding that the task was beyond the test-taker's level of language ability.

In the writing task, the test-taker's emotional regulation processes used to cope with negative emotions were task-focused: he focused on the letter format and on the presence of the prompt in both Arabic and English. In the beginning of the listening and grammar sections the test-taker's confidence was high as reflected by phrases he used such as: "it looks easy, beautiful exam like for children." Here there was no indication of regulation of emotions. The test-taker focused on the task and expressed impatience in the listening section at the beginning i.e. he focused on the task. When the test-taker became anxious, the coping processes used were feeling angry and blaming the test designers for making the test difficult (emotion-focused) and at the same time feeling that the exam was a challenge that would make students study more (task-focused). The test-taker alternated between task-focused and emotion-focused processes. Here it is very difficult to separate the emotions from the emotion-focused process. The test-taker felt angry at the test designers. One of the test-taking strategies used by the test-taker was translation for items he got both correct and incorrect and thus, it is difficult to decide whether it is a contributory or non-contributory strategy. The non-contributory strategies he used were *guessing, formulating ideas in Arabic and then translating into English, using words from the prompt and reading aloud a possible answer and then changing and selecting a final answer*. In attempting to identify the meaning of a word, the test-taker used two non-contributory strategies: *pronouncing the target word* and using the *context to guess the meaning*.

The behavioral outcomes of a test-taker's utilization of test-taking strategies and emotional regulation processes can be persistence, increase or decrease of effort or giving up. The main mode of behavior of Test-taker A was to give up when anxiety was high. This occurred in the listening, grammar and reading sections.

Overall, the LTP does explain Test-taker A's test-taking process with four discrepancies: it is difficult to classify translation as either a contributory or non-contributory strategy, coping processes used included task-focusing processes and it is difficult to identify processes that regulated emotions when the test-taker experienced positive emotions. It was also difficult to separate emotions from emotion-focusing processes.

5.2.2 Test-taker B (Female, elementary English language proficiency stage, 18-22 age range)

5.2.2.1 Writing

The test-taker *read the prompt in English and translated as she read*. She found some words to be difficult but stated that the topic was easy. She thought it was a good idea to have the prompt in Arabic but she would have preferred to have only the difficult words translated and not the whole prompt. She stated that in this way students would be independent and rely on themselves. The test-taker expressed concern that “if foreigners saw these materials in Arabic, then they will think that we are making it too easy as we do not know or cannot understand English.” The test-taker then *read the prompt in Arabic*.

She negatively assessed her ability to perform the task. She stated that writing a letter was difficult for her because had not learned how to write letters and she did not know what to write. She did not have enough words in English to explain the relationship between herself and colleagues and between a boss and subordinates. She felt that this task would be difficult because a letter to a boss should have a particular structure and organization which she was not familiar with.

Because she did not know enough words in English she stated that her strategy was to *use some words in the prompt*. She did not approve of this strategy because this is what happens in the preparatory and secondary school exams where students do not know any other words to use. She then proceeded to write the letter. She would *think of an idea / sentence in Arabic, translate it to English and then write it down*. She did this both at the word and sentence levels. She stated twice that she was not sure of what she had written. The test-taker wrote a paragraph dominated by structural errors and run-on sentences and it was difficult to understand what she wrote.

5.2.2.2 Listening

In Part A the test-taker expressed approval for including a listening section in the exam. She lost attention or concentration twice and got 7 items correct. In Part B she had more difficulties and got only 4 items correct. She started to find some items difficult to understand or to hear stating that they were too fast. In Part C she faced more

difficulties and she started to blame the difficulties faced on the tape quality. She stated that the quality of the tape is not good. Then she stated that “for a beginner like me, 10 dialogs are too much” and “it is too hard for me to answer 10 spoken dialogues.” (Here her tone sounded as if she was saying it’s not fair). She left 6 items blank in this section because she found them too difficult or spoken too fast or some of the words used were unfamiliar. The strategies she used were *paraphrasing into Arabic and guessing*.

5.2.2.3 Grammar

The test-taker started out with a negative self appraisal of ability stating “I know very little about grammar as I graduated from a commercial school and not a high school.” She *read the instructions in Arabic* and stated that she was familiar with the multiple-choice format of questions. She found the *example item* easy and hoped that the rest of the items would be as easy. She found the first few items to be easy but gradually getting more difficult because of her not knowing the meaning of some of the words. The main strategy she used was *translation*. She also used *background knowledge: syntax and lexical* (Examples: “I know that ‘-ing’ is used to indicate continuity.” “The answer is shouldn’t because he is very nervous.”) and *insertion of each option one at a time in the stem*.

She found Part B very difficult (she stated this three times) because she had never dealt with this type of question before. She used *guessing, translation* and *she ran out of time without completing all the questions*. The test-taker got 7 items correct (out of 15) in Part A and 2 items correct (out of 15) in Part B.

5.2.2.4 Reading

The test-taker started this section of the test with an initial feeling of relief stating that she was familiar with the format of a reading comprehension test. The test-taker attempted 3 passages. She read the first passage and *translated* as she read but found some difficult words. She then tried to answer the questions but was not successful. She persisted and read the second passage and correctly identified the gist. She *correctly guessed the meaning of ‘guarantee’ from the overall meaning*. She found

passage three to be full of “heavy” and difficult words and *decided not to continue*. She got only 1 item correct in this section.

5.2.2.5 Affective Responses

The following Table (18) illustrates the emotions that emerged from the think aloud protocol of Test-taker B as well as the indicators of these emotions.

Table (18): Affective responses of Test-taker B

Emotion	Indicators from the Think Aloud Data
Disapproval	She did not want the whole prompt to be translated into Arabic, only the difficult words. The prompt in Arabic to her reflected that the test was for low ability test-takers.
Anxiety	She was anxious about writing a letter and about the particular writing task stating that it was difficult for her. She experienced anxiety in each of the listening, grammar and reading sections: Listening: She stated that the task was too hard for her. Grammar: She stated she knew very little grammar. In part B she repeated several times that the questions were difficult. Reading: She stated that passage three was too difficult for her and she was not going to continue. She handled anxiety by: persisting and guessing, translating from English to Arabic or translating from Arabic to English (writing), giving up and blaming the poor quality of the tape in the listening task.
Anger	Blamed the poor quality of the tape in the listening task (caused by anxiety).
Approval	Initially the test-taker started out expressing approval about a listening test but gradually faced difficulties and started to complain about the quality of the tape and difficulty of the task.
Relief	At the beginning of the reading section, the test-taker expressed relief because she was familiar with the format of a reading comprehension test. This feeling of relief changed to anxiety and then she gave up.
Confusion	She could not understand an item that was easy in the listening section.
Frustration	Here frustration at the speed of the tape in the listening section was reflected from her tone.

5.2.2.6 Fit of the LTP Model

At the beginning of each test task: writing, listening grammar and reading tasks, the test-taker assessed her language ability to perform these tasks. The metacognitive strategies used were: negative assessment of language ability in a specific skill or task (the test-taker did not feel confident about her ability to write letters), negative

assessment of language ability in a specific skill attributed to educational background, negatively assessing a specific feature in the design of test task (the test-taker did not approve including a full translation of the writing prompt), positively assessing a specific feature in the design of test task (the test-taker approved the inclusion of a listening task in the test), predicting that the task would be easy because of familiarity with the test task format and deciding that the task was beyond the test-taker's level of language ability.

In the writing task, the test-taker's coping process focused on the difficulty of the task compared to her ability. In the very beginning of the listening section before the test-taker reported experiencing any negative emotions, the test-taker commented that: "this listening is very good, this is easy" (once again these are task-focusing processes but there was no indication of regulation of emotions) and she then blamed the poor quality of the tape in the listening task (emotion-focused). In the beginning of the grammar section, the test-taker was neutral hoping that that the rest of the items would be as easy as the example item (focused on the task). The main test-taking strategy used by the test-taker was translation for items she got both correct and incorrect and it is difficult to decide whether it is a contributory or non-contributory strategy. Contributory strategies used were *background knowledge: syntax and lexical and guessing the meaning of a word from the context* (this was a non-contributory strategy for Test-taker A). The main non-contributory strategies used in sections where she performed poorly (writing, grammar and reading) were *formulating ideas in Arabic and then translating into English, using words from the prompt, guessing and insertion of each option one at a time in the stem*. The main mode of behavior of Test-taker B was to give up when anxiety was high.

Overall, the LTP does explain Test-taker B's test-taking process with two discrepancies: both contributory and non-contributory strategies were used when there was a negative affective response and it was difficult to separate emotion from the emotion-focusing process. The model postulates that only non-contributory strategies are utilized when a test-taker experiences negative emotions.

5.2.3 Test-taker C (Male, elementary English language proficiency stage, 23+ age range)

5.2.3.1 Writing

The test-taker started to read the prompt in English and initially did not understand. He persisted and continued to read stating that maybe the context would help him understand the topic better. He read the prompt in English and translated into Arabic as he read. He stated that he understood the meaning and that it became clearer as he read. He then read the prompt in Arabic to make sure he understood. Similar to Test-takers A & B, he had doubts about his ability to complete the task and stated that writing a letter “may be above my current level but God willing, I will do it.” He stated that the provision of an Arabic prompt helped weak students.

He described the strategy he used as *‘squeezing his brain’ for expressions* and stated that if he were not taking an English course at the time he would not be able to address the topic and words will not ‘jump’ out for him to write immediately. He believed that *reading the prompt in English and translating into Arabic* helped completion of the task. The test-taker wrote a paragraph of one run-on sentence dominated by structural errors and reflecting a lack of knowledge of sentence boundaries and structure. It was difficult to understand what he wrote.

5.2.3.2 Listening

Test-taker started by describing his previous experience of taking a placement test and apparently it was a negative experience. He stated that there was no listening task and all he “did was writing sentences and filling in exercises.” When he received the answer sheet (a standard computerized answer sheet with bubbles that is scored by an optical scanner) he found it rather scary.

He found the speaker to be fast but was able to *follow the instructions by reading along*. Test-taker C’s experience was similar to that of Test-taker B. In Part A he initially he found the items to be very easy. He expressing approval stating that it was a good exam, comprehensive and the pauses between the items were sufficient for him to think and answer. Towards the end of Part A, he stated that the test was becoming more difficult. He was *not sure of the previous two items*. In Part B the main strategies used

were *reading the options and translating*. He still considered the pauses between the items to be sufficient, however, he started to complain that the exam is too long and that some items are difficult. By item 27 he started to blame the quality of the tape and that the time is not sufficient to think. He stated that the tape should be at a higher volume, he could not understand what was said and the voice was not clear. He ended by stating that his answers were not wrong but the tape was not clear. In Part C, the process was similar. He *read, translated the options* and used *guessing*, complaining about the difficulty of the items and the tape not being clear. He got 9 items correct in Part A, 9 items correct in Part B, and 3 items correct in Part C.

5.2.3.3 Grammar

He negatively assessed his grammar ability by saying that he was not good at grammar. He was very pleased to find an example for each item by stating: “the beauty of this is that there is an example for each item.” By item 48, the test-taker he stated that he was feeling bored and that the exam was getting too long. The main strategy used was *reading the item and options*. He found Part B to be very difficult because he did not know the meaning of some of the words. He wished there could be a translation but he knew that was impossible (wishful thinking). The main strategy used was just *guessing*. He got 12 items correct (out of 30) in this section.

5.2.3.4 Reading

The test-taker *read the instructions in Arabic*. He read the first passage and found a lot of difficult words but continued reading. He then *read the items and options* and *reread parts of the passage* to answer the questions. He stated that “all is difficult, I cannot understand anything, all difficult.” He then moved to the second passage. At this point he was worried about the length of the second passage stating “all this in one passage, God help me”. However, after reading it, he found it easier than the first one. *He read the items and options*. The test-taker then *gave up* stating that it took him half an hour to get through 2 passages only. He got 3 items correct in this section.

5.2.3.5 Affective Responses

The following Table (19) illustrates the emotions that emerged from the think aloud protocol of Test-taker C as well as the indicators of these emotions.

Table (19): Affective responses of Test-taker C

Emotion	Indicators from the Think Aloud Data
Approval	Initially the test-taker started out expressing approval about a listening test but gradually faced difficulties and started to complain about the quality of the tape and difficulty of the task.
Pleasure	He was pleased to find example items in the grammar section.
Anxiety	He experienced anxiety in each of the listening, grammar and reading sections: Listening: As from question 14 he repeated several times that “it is difficult”, “I can’t undertand”. Grammar: He stated he was not good at grammar and by question 48 he stated that he was bored and the exam was too long. Reading: He repeated several times that “all is difficult” and that he could not understand anything. He handled anxiety by: persisting and guessing (MCQ), translating from English to Arabic (MCQ), or translating (writing), giving up and blaming the poor quality of the tape in the listening task.
Boredom	He stated that he was bored when he found the items in the grammar section to be very difficult
Anger	Blamed the poor quality of the tape in the listening task

5.2.3.6 Fit of the LTP Model

At the beginning of each test task: writing, listening grammar and reading tasks, the test-taker assessed his language ability to perform these tasks. The metacognitive strategies used were: negative assessment of language ability in a specific skill or task (the test-taker did not feel confident about his ability to write a letter), comparing current test situation with a previous test-taking experience, positively assessing a specific feature in the design of the test task (the test-taker approved of the listening task and of the Arabic instructions), predicting that the task would be difficult because of length (too long) and deciding that the task was beyond the test-taker’s level of language ability.

In the very beginning of the listening section the test-taker’s assessment included recalling his negative emotions experienced during a previous placement test. In the writing task the test-taker’s coping processes used were task-focusing: he focused on the presence of the prompt in both Arabic and English. In the first part of the listening

section experienced positive emotions as reflected by phrases such as: “I think it is easy”, “This is a good exam.” When negative emotions emerged, the test-taker’s coping processes used were emotion-focusing and task-focusing. The emotion-focusing processes were feeling bored, wishful thinking and external attribution of blame: the poor quality of the tape stating that “the tape should be louder” and “voice is not clear” and that the exam was too long. Once again these emotion-focusing processes are actually emotions. The task-focused coping process used was focusing on a particular aspect of the test: “the beauty of this is that there is an example for each item.” The main test-taking strategy used by the test-taker was translation and similar to Test-taker B it is difficult to decide whether it is a contributory or non-contributory strategy. The main non-contributory strategy used was *guessing*. When negative emotions emerged, Test-taker C persisted in the writing, listening and grammar tasks. However, in the reading task he gave up.

The fit of the LTP model is the same as that of Test-taker B’s test-taking process with three discrepancies: not all test-taking strategies can be classified as either contributory or non-contributory, when there is a negative affective response coping processes included both task-focusing and emotion-focusing processes and emotions are difficult to separate from emotion-focusing processes. The model postulates that only emotion-focusing processes are utilized when a test-taker experiences negative emotions. It is difficult to identify emotional regulation processes when positive emotions were experienced.

5.2.4 Test-taker D (Female, elementary English language proficiency stage, 23+ age range)

5.2.4.1 Writing

The test-taker *read the prompt in English and translated it*. She then *read the Arabic translation* provided. She emphasized that she translated the prompt before looking at the Arabic translation. She wrote the greeting and introduction of the letter. She *translated into Arabic what she wrote in English*. She *used words from the prompt in English* when she could not find the English words to express what she wanted to say. She continued to write and translate or *formulate her ideas in Arabic and translate them*

into English. No affective responses were mentioned. The test-taker wrote two paragraphs consisting of run-on sentences and fragments dominated by structural errors and reflecting a lack of knowledge of sentence boundaries and structure. There was some communication of ideas.

5.2.4.2 Listening

In Part A she *left 4 items blank* because she did not hear or did not understand what was said. She found the last 3 items: 13, 14 & 15 to be easy and she got them correct. She started to feel worried and she stated that she was not looking forward to the next section (negative appraisal). In Part B she found several items difficult to understand and by item 27 she started to complain that the listening task was too fast and too difficult. She stated: “too fast, one cannot think. At first, it was slow, now it is too fast.” She stated that she could understand the individual words but not the overall meaning of what was said. In Part C she found the items difficult and did not expect that there would be 2 utterances in each exchange. At item 37, she complained that there was something wrong with the cassette recorder. The main strategy used was *translating the item (what she heard) into Arabic*. She left *11 items blank* in Parts B & C. She got 7 items, 3 items and 3 items correct in Parts A, B, & C respectively.

5.2.4.3 Grammar

The test-taker stated that “writing instructions in Arabic is a good idea because it helps beginner students understand the task, thus, it gives more accurate measurements and at the same time it does not hinder the advanced students.” The main strategies used were *reading the instructions, reading the item and options, and translating*. At the end of Part A she expressed surprise that the last item was easy after she felt that the questions were getting tougher. She thought that in a test the items should be more difficult at the end. She got 8 items correct in Part A. In Part B she read the instructions and translated them. Then she *read each item and translated*. She got only 1 item correct in Part B. She did not report any affective responses.

5.2.4.4 Reading

She started by a negative appraisal saying that her problem was with vocabulary. She attempted 3 passages only and she *translated into Arabic as she read*. The strategies she used were *reading the item and translating and eliminating options*. She got 4 items correct.

5.2.4.5 Affective Responses

The following Table (20) illustrates the emotions that emerged from the think aloud protocol of Test-taker D as well as the indicators of these emotions.

Table (20): Affective responses of Test-taker D

Emotion	Indicators from the Think Aloud Data
Anxiety	She experienced anxiety in each of the listening and reading sections: Listening: At the beginning of part B she stated she was not looking forward to the next section and the repeated several times throughout “it is too difficult” and “I did not understand”. Reading: She started the section by stating that she had a problem with vocabulary. She handled anxiety by: persisting and guessing, translating into Arabic, leaving items blank, giving up (in the reading) and blaming the poor quality of the tape in the listening task.
Approval	She expressed approval about having the instructions in Arabic.
Anger	Blamed the cassette recorder for not working well
Frustration	She expressed her frustration several times about the speed of the tape in the listening section.
Confusion	She stated that she was confused in an item in the listening section.

5.2.4.6 Fit of the LTP Model

At the beginning of Part B of the listening task and the reading task, the test-taker assessed her language ability to perform these tasks. The metacognitive strategies used were: negative assessment of language ability in a specific skill or task (the test-taker did not feel confident about her ability to handle Part B of the listening task), and predicting that items would be difficult towards the end of a section.

In the first part of the listening section the test-taker did not experience negative emotions as demonstrated by phrases such as: “This is easy” and “Not bad.” In the

beginning of the grammar section the test-taker did not experience negative emotions and focused on the Arabic instructions (it would be difficult to conclude here that focusing on the instructions is a process used to regulate emotions). When negative emotions emerged, the test-taker's emotion-focused coping processes used were blaming the cassette recorder for not working well and the tape for being too fast. The main test-taking strategy used by the test-taker was translation and once again it is difficult to decide whether it is a contributory or non-contributory strategy. A contributory test-taking strategy used was eliminating options. The non-contributory strategies she used in the writing section where she performed poorly were *formulating ideas in Arabic and then translating into English and using words from the prompt*. As with Test-taker C, when negative emotions emerged, Test-taker D persisted in the writing, listening and grammar tasks. However, in the reading tasks he gave up.

The LTP model does fit Test-taker D's self-reported test-taking process with three discrepancies: translation (the main test-taking strategy used) can not be classified as either contributory or non-contributory, it is difficult to identify emotional regulation processes when positive emotions emerged and it is difficult to separate emotions from emotional-regulation processes.

5.2.5 Test-taker E (Male, intermediate English language proficiency stage, 18-22 age range)

5.2.5.1 Writing

The test-taker *read the prompt in English and translated as he read*. He found many words to be difficult and this was made easier by the Arabic translation of the prompt. He understood the task assigned which was to write a letter explaining the error and its consequences. He found the time allocated to be sufficient however, he did not like the task. He did not find it appropriate, relevant or interesting. He stated "What does it have to do with me to write a letter to Xerox?" He preferred to write on a general topic related to pollution, overpopulation, ... that all students can handle. The test-taker felt that a business major might have an advantage and handle the letter better than he would since the test-taker may have vocabulary in another area. He thought that the instructions indicating that the "letter layout is not part of the grade may be overlooked

by many students, especially those who are nervous.” The test-taker also stated that the directions were not clear regarding how long the letter should be, whether 10 sentences or 10 lines. The test-taker *formulated his ideas in Arabic and then wrote them down in English*. He commented once again that credit or grades should be given for the correct letter format. The test-taker wrote a paragraph that consisted of short sentences that were grammatically correct. He *wrote one sentence from the prompt* and he did not complete the last sentence. There was little communication of his own ideas.

5.2.5.2 Listening

In Part A the test-taker did well in this part and got 11 items (out of 15) correct. He believed that using pictures was a good idea. In Part B he had more difficulties and got only 7 items correct. By item 22 he stated that the questions were getting more difficult and by item 28 the test-taker complained that the test was too long and he could not concentrate. In Part C he got 6 items correct (out of 10). However, at the end he thought that 40 items for listening were too many and he complained that he had a headache and was unable to concentrate. He stated that listening to the utterance once only was a problem because a test-taker may be tired and thus, does not hear clearly. Therefore, according to the test-taker the solution is to repeat the sentences or mini-dialogues twice.

5.2.5.3 Grammar

The test-taker *read the instructions and example*. The strategies used were *reading the items and substituting each of the distracters into the stem to choose the right answer*. In Part A he stated he considered this section of the test to be the easiest. He explained that this is because “We all study this. We all review grammar before the exam.” The test-taker found Part B to be difficult and by item 59 he stated that “This section is as hard as the listening.” He *monitored his time* and twice mentioned that he needed to answer faster. The test-taker got 9 items correct (out of 15) in Part A and 4 items (out of 15) correct in Part B.

5.2.5.4 Reading

The test-taker started this section of the test with a negative appraisal stating that it was too long and needed time. The main strategy used is to *look at the question and then look at the passage to locate the line where the answer to the question is*. He *guessed* the answer for the item on vocabulary in context because he did not know the meaning of any of the options. He did not like this type of item. He got 12 items (out of 20) correct in this section.

5.2.5.5 Affective Responses

The following Table (21) illustrates the emotions that emerged from the think aloud protocol of Test-taker E as well as the indicators of these emotions.

Table (21): Affective responses of Test-taker E

Emotion	Indicators from the Think Aloud Data
Disapproval	He did not like the writing task and thought it was unfair. He considered the instructions and directions to be inadequate. He thought that points should be given for correct letter format. In the listening, he thought that the statements and dialogs should be repeated twice.
Anxiety	He experienced anxiety in each of the listening, grammar and reading sections: Listening: By question 28 he stated that “it is long... I can’t concentrate”. He stated that he had a headache. Grammar: In part B he stated that it “is as hard as the listening”. Reading: He started by stating that it was too long and needed time. He handled anxiety by: persisting and guessing, and complaining of a headache and not being able to concentrate.
Approval	He expressed approval about the use of pictures in the listening task.
Complaining	He complained that the listening and reading tasks were too long.
Frustration	He expressed frustration with the vocabulary items in the reading section.
Confusion	He stated that some items had two answers in the listening section.
Tiredness	He stated that he felt tired at the end of the listening section.

5.2.5.6 Fit of the LTP Model

At the beginning of the writing and reading tasks, the test-taker assessed his language ability to perform these tasks and in both there was a negative affective response. The metacognitive strategies used were: negative assessment of language

ability in a specific skill or task (the test-taker did not feel confident about her ability to write letters), negatively assessing a specific feature in the design of the test task (the test-taker did not approve of the writing task and having the utterances in the listening task spoken only once), positively assessing a specific feature in the design of test task (the test-taker approved of using pictures in the listening task), predicting that the task would be difficult because of length (too many reading comprehension passages) and deciding that the task was beyond the test-taker's level of language ability.

In the writing task the test-taker's coping process used was emotion-focusing: he complained about and disapproved of the task (these are actually emotions). In the listening and grammar sections the positive emotions experienced by the test-taker were reflected by phrases such as: "questions are graded in difficulty", "pictures help a lot", and "so far, sentences are easy." When negative emotions emerged, the test-taker's coping processes used were emotion-focusing and task-focusing processes. The emotion-focusing processes were feeling tired, losing concentration and feeling that there were too many items and reading passages (once again these are all emotions). The task-focused coping process used was monitoring time. It was difficult to classify the test-taking strategies used by the test-taker as either contributory or non-contributory. In the above analysis of Test-takers A, B, C & D's test-taking processes, the researcher identified the following non-contributory strategies used in the writing section where they performed poorly: *formulating ideas in Arabic and then translating into English and using words from the prompt*. Test-taker E who is at a higher proficiency level and performed better in the writing task used the same strategies as Test-takers A, B, C & D. Other strategies used that were difficult to classify as either contributory or non-contributory were *reading the instructions* and *monitoring time*. Two strategies that were both contributory and non-contributory were *substituting each of the distracters into the stem to choose the right answer* and *looking at the question and then looking at the passage to locate the line where the answer to the question is*. The main non-contributory strategy used was *guessing*.

When negative emotions emerged, Test-taker E persisted in all test tasks. The fit of the LTP model is the same as that of Test-taker C.

5.2.6 Test-taker F (Female, intermediate English language proficiency stage, 18-22 age range)

5.2.6.1 Writing

The test-taker *read the prompt in English and then read the prompt in Arabic*. She was pleased that the prompt was translated into Arabic and thought that it was a good topic. She stated that the prompt did not specify how many sentences were expected. She was not happy that the directions indicated there were no points given to the layout of the letter. She stated that if test-takers know that they will gain some points for the layout, this will reduce their anxiety and help them perform better. She added that “We are used to this in the public school exams.”

She felt unable to start to write and decided to *use a lot of words from the prompt*. She then mentioned marks or points for the layout of the letter once again. “If you make things easier and give us the information about to whom the letter will be sent, the address and the signature then we could have obtained a higher score.” The test-taker wrote a paragraph that consisted of only three sentences with some grammatical errors using words from the prompt.

5.2.6.2 Listening

In Part A she was pleased that the directions were translated into Arabic. The test-taker had difficulties in this part and got 7 items (out of 15) correct. She *guessed* the answer for several items and got them correct. Her tone sounded angry when she stated that the speaker on the tape must repeat the sentence. She said that each sentence should be said twice. She mentioned this 3 times. In Part B she was worried and started by stating once again that “usually a test-taker is very nervous and that may affect the ability to hear so the listening should be repeated.” She got 7 items correct in this part. In Part C the test-taker got 3 items correct (out of 10) in this part. At the beginning of this part she stated that the sound was not clear and the utterances should be repeated twice. She was not able to hear / comprehend 7 items and noted that she had to *guess* a lot.

**PAGE
MISSING
IN
ORIGINAL**

Table (22): Affective responses of Test-taker F (continued)

Emotion	Indicators from the Think Aloud Data
Anxiety	She experienced anxiety in each of the listening, grammar and reading sections: Listening: She repeated many times that she couldn't hear and utterances should be repeated twice. She stated several times that she had to guess. Grammar: In part B she stated several times that it was confusing. Reading: She stated that the reading took time. She stated that the questions were difficult and that she had to guess. She handled anxiety by: persisting and guessing, and giving up.
Anger	She expressed anger because she thought she could have scored a higher score in writing if the layout of the letter had been taken into consideration. She also expressed anger about the listening tasks being heard only once.
Boredom	She expressed boredom in the reading section.
Frustration	She expressed frustration at the speed of the tape in the listening section and lack of time in the reading section. She wanted it made shorter and she stated she was really fed up.
Confusion	She stated she was confused in several items in the grammar section.

5.2.6.6 Fit of the LTP Model

Similar to the previous test-takers' profiles, before each test task Test-taker F assessed her ability to perform the task resulting in either a positive or negative emotional response. The metacognitive strategies used were: negative assessment of language ability in a specific skill or task (the test-taker found it difficult to start to write), negatively assessing a specific feature in the design of test task (the test-taker did not approve of not giving points for layout of the letter), positively assessing a specific feature in the design of test task (the test-taker approved of having the prompt and directions in Arabic), predicting that the task would be easy because of familiarity with the test task format (section A of the grammar task), predicting that the task would be difficult because of test task format (section B of the grammar task) and deciding that the task was beyond the test-taker's level of language ability.

In order to complete each task, the test-taker utilized a combination of emotional regulation processes and cognitive test-taking strategies. In the writing task the test-taker's coping processes used were both task-focusing and emotion-focusing: she

focused on the presence of an Arabic prompt and complained (emotion) about not being given points for the layout of a letter. In the listening and grammar sections the test-taker experienced positive emotions as were reflected by phrases such as: “I heard this well” and “this is a good item” and focus on the instructions being in Arabic. When negative emotions emerged, the test-taker’s coped by using emotion-focusing processes. In the listening task, the emotion-focusing processes were blaming the voice on the tape for not being clear, and complaining about not having enough time to respond and that the utterances should be repeated twice (emotions). Other emotions expressed were feeling confused, repeating “I don’t know” and feeling fed up. Similar to Test-taker E, it was difficult to classify the test-taking strategies used by the test-taker as either contributory or non-contributory. For example, in the reading task the strategies of *reading the passage first and then answering the questions and starting by reading the questions first and then the passage* led to the same results. She only got one question correct on each passage. For Test-taker F, guessing was both a contributory and non-contributory strategy. She correctly guessed the answer to several items in the listening task.

The outcome of the processes used when negative emotions emerged was persistence in the writing, listening and grammar tasks. However, she gave up in the reading task. The fit of the LTP model is the same as that of Test-taker C.

5.2.7 Test-taker G (Male, intermediate English language proficiency stage, 23+ age range)

5.2.7.1 Writing

The test-taker started by a negative appraisal of his ability to handle the task. He stated that the topic was completely new and the problem was that he had never written a letter to a company. He did not have the vocabulary to address a general manager and he is not used to writing to one. His tone sounded agitated and he stated that he felt nervous because he lacked the language to express the intended meaning and felt very uncertain about choosing the right words. He stated that “I am at a loss on what to write. A paragraph would have been better.” The test-taker then wrote. He did not mention

any of the strategies he used. He has knowledge of sentence boundaries, however, his writing was dominated by spelling and grammatical mistakes.

5.2.7.2 Listening

In Part A the test-taker started with high anxiety stating that the items require concentration and that this might be very difficult for the beginner. However, he gained confidence and by item 12 stated that “this is for children, like children stories.” He reported that the main strategy he used was to *rely on grammar*. At the beginning of Part B he suggested that to save time, the directions are given in Arabic only. He hoped that the questions would be easy and that he would be able to answer them all. He found this part to be more difficult and by item 25 he started to get tired after getting 4 items in a row wrong. He stated “thank god we are about to finish.” He got 12 items (out of 15) correct in Part A and 8 items (out of 15) correct in Part B. In Part C the test-taker performed much better and got 8 items correct (out of 10). At the beginning of this part he requested that the utterances are said twice rather than one time. In the second item he started to gain confidence and stated that “this is better”. In the last item he commented that he thought there was an opportunity for cheating (perhaps wishful thinking).

5.2.7.3 Grammar

In spite of the fact that the test-taker was negative about his ability in grammar stating he didn't study grammar well and had he reviewed he could have answered all the items, he performed well in Part A (he scored 13 items out of 15 correct). Although he considered the questions to be good, yet many of them he thought did not test grammar but tested special difficulties. He thought that grammar questions should be focused on tenses and not on *meaning of words* such as should or shouldn't. The main strategy used was *reading the item and options*. The test-taker found Part B to be difficult. He stated that he felt nervous in this part. He found the sentences to be OK since he did not know the rules being violated. The mistakes were not noticeable and he felt that there was not enough time. He stated that 30 seconds are not enough and this made him more nervous. It made him lose concentration to know that when he reads the

item he has less than a minute. He *read the difficult sentences over and over again and spent too much time on them*. He got 8 items correct (out of 15) in Part B.

5.2.7.4 Reading

From the very beginning the test-taker was negative and asked, “are we back to Thanawiya Amma?” (high school). He *read the questions first*. In the first passage, he stated that there is a trick in question one. He believed that items are tricky and the questions are not simple. He found the exam to be too long with too many reading comprehension passages. He compared this placement exam with the old one and said that the present one was better in terms of item type. He found the questions about what the writer means by this word or expression to be difficult and *often just guessed the answer*. He *monitored time passing and tried to look for the answers in the fastest way possible*. He got 12 (out of 20) items correct in this section.

5.2.7.5 Affective Responses

The following Table (23) illustrates the emotions that emerged from the think aloud protocol of Test-taker G as well as the indicators of these emotions.

Table (23): Affective responses of Test-taker G

Emotion	Indicators from the Think Aloud Data
Approval	He expressed approval about the item types included in this test compared to the previous one.
Disapproval	He expressed disapproval about the writing task since he had never written a letter to a company before.
Anxiety	He experienced anxiety in all the sections although he performed well in the listening section and Part A of the grammar section. Writing: He stated that he felt nervous about the task. Listening: He stated several times he did not hear and he had to guess several items in a row. By question 25 he stated that he was getting tired of the exam. Grammar: He stated that he felt nervous in Part B. Reading: He stated that the passage was getting more difficult. He expressed concern about time passing and needing time to answer. He guessed the answer to several questions. He handled anxiety by: persisting and rereading or guessing.

Table (23): Affective responses of Test-taker G (continued)

Emotion	Indicators from the Think Aloud Data
Wishful thinking	He thought that there could be an opportunity for cheating although he had done well in that section (Part C of the listening task).
Complaining	He complained that the reading tasks were too long.
Frustration	He expressed his frustration at the lack of time in the listening, grammar and reading sections. He mentioned time all throughout.
Confusion	He stated that he was confused in several items in the grammar section.
Tiredness	He that he was tired during the listening section.
Relief	He stated that he felt relieved that the listening section was about to finish.

5.2.7.6 Fit of the LTP Model

The description of the fit of the LTP model is similar to that the previous test-takers' profiles with no new patterns emerging from the data except for the following metacognitive strategy used by Test-taker G: comparing the current test design to another test.

5.2.8 Test-taker H (Female, intermediate English language proficiency stage, 23+ age range)

5.2.8.1 Writing

The test-taker *read the prompt in English and then in Arabic*. The test-taker was confident and stated that one can write using a lot of vocabulary and as much knowledge as one knows. She *formulated her ideas in English* and wrote them down. She stated that the writing task is very good because “the topic is written in English and the Arabic translation helps the intermediate student. We may forget some words.” “It gives a chance for all to understand the writing prompt.” She thought that thirty minutes was too long and she needed only 15 minutes to write the letter. She wrote 9 short sentences containing both grammatical mistakes and some lexical errors. The last sentence was not even comprehensible: “I want to stand of my work in this department.”

She stated that she did not like multiple choice questions at all. (This was clearly shown in the MCQ component of the test. She performed very poorly in all sections: listening, grammar and reading).

5.2.8.2 Listening

The test-taker started by stating that she did not understand anything. Then she was pleased (she stated 'good') because there was an Arabic translation of the directions. She did not understand or did not hear several items. In 3 of the 4 items where she commented 'good' or 'very good' she gave the wrong answer. In Part A she got 4 items (out of 15) correct. In Part B her performance was slightly better (she got 5 item out of 15 correct), yet she felt it was more difficult. By item 23 she started to complain that the items were difficult. In Part C her anxiety was high and she stated that some of the questions were too difficult which made her feel helpless. The strategies used were *reading options and guessing*. The test-taker got 3 items (out of 10) correct. At the end of the listening test she stated: "I can't hear. Even if I know something, I can't hear it. It makes me nervous and tense when I can't hear the question and then I guess the answer. As a result, I don't hear the next question. In this way, the result will be inaccurate. It will give an inaccurate assessment of the proficiency level of the student."

5.2.8.2 Grammar

The test-taker started to *read the directions in English but then switched to the Arabic translation*. The test-taker got 6 items correct (out of 15) in Part A and 1 item correct (out of 15) in Part B. In spite of the fact that the test-taker did poorly in Part A, yet she stated that this section was excellent. The reason she gave was as follows: "If I couldn't answer some items, it is because I don't know the answer, not because I can't hear." She attributed her difficulties in the listening task to problems in hearing and not in comprehending the meaning. The test-taker did not seem to be aware that she was not doing well. The main strategies used were *reading the item and options and translating*. The test-taker found Part B to be extremely difficult. She *read the directions for this part twice in Arabic* to understand the task. She attempted only 5 items (only 1 was correct) and then *gave up*. She found this part of the test to be too difficult.

5.2.8.4 Reading

From the very beginning the test-taker was negative and found passage one to be difficult. She *read and reread passages one and two and the questions and options. She read portions of the passage to answer questions and she translated.* She read passage three and stated that she only understood the general meaning but not the details. She could not answer and *gave up* stating that it required a high level of language ability. She thought it was a good idea to include reading comprehension in the test, however she needed more time. She got 2 (out of 20) items correct in this section.

5.2.8.5 Affective Responses

The following Table (24) illustrates the emotions that emerged from the think aloud protocol of Test-taker H as well as the indicators of these emotions.

Table (24): Affective responses of Test-taker H

Emotion	Indicators from the Think Aloud Data
Approval	She expressed approval about the writing task and the prompt being provided in both Arabic and English. She approved of the grammar and reading tasks.
Disapproval	She expressed disapproval of MCQ tasks which was reflected in her performance.
Anxiety	She experienced anxiety in each of the listening, grammar and reading sections: Listening: She started Part C by stating that some questions were too difficult and made her feel it was hopeless. Grammar: In Part B she repeated several times that it was difficult and by question 61 she decided not to continue because it was too difficult and it needed more time. Reading: In the third passage she stated that she could only understanding the general meaning and not the details and so she was unable to answer. She stated that it required a higher level of language ability. She handled anxiety by: persisting and rereading, translating, guessing or giving up.

Table (24): Affective responses of Test-taker H (continued)

Emotion	Indicators from the Think Aloud Data
Helplessness	She stated that she felt helpless in the listening section, particularly in Part C. She stated that this part of the test would give an inaccurate assessment of her proficiency level.
Frustration	She expresses frustration at the lack of time in the grammar section and decided not to complete it.
Confusion	She stated that there were two correct answers in an item in the listening section and she did not understand what she should do in the first passage in the reading section.

5.2.8.6 Fit of the LTP Model

The description of the fit of the LTP model is similar to that of the previous test-takers' profiles with no new patterns emerging from these data.

5.2.9 Test-taker I (Male, advanced English language proficiency stage, 18-22 age range)

5.2.9.1 Writing

The test-taker *read the prompt in English and then summarized the writing task in English*. He then *formulated his ideas in Arabic*. He noted that this task is more suitable for business and it would be easier to write on a social topic since he could write more. He wrote fluently in short sentences with few grammatical errors. There was good communication of ideas.

5.2.9.2 Listening

In Part A the test-taker was confident and *translated some of the utterances*. He did well got 11 items (out of 15) correct. In Part B he did better and got 13 items (out of 15) correct. He found the pause between the items to be sufficient but did not know or did not hear 4 items where he had to *guess the answer* (he guessed 3 of them correctly). In Part C he got 8 items correct (out of 10). Although he did well she found this part to be difficult and confusing. After he listened to the mini-dialog he *read the options* and found that he forgot what was said.

5.2.9.3 Grammar

The test-taker got 9 items correct (out of 15) in Part A and 12 items correct (out of 15) in Part B. Although he commented that in Part B some items were confusing, he did much better than in Part A. He also got those items he thought confusing correct.

5.2.9.4 Reading

The test-taker started by being negative stating that this looks difficult. He got 11 items (out of 20) correct in this section. By the end of Passage 1, he was yawning a lot. By the third passage, he noted that it was difficult and it was taking a long time. He stated that she had to *reread either parts of the passage or the whole passage each time to answer a question*. He found the last passage to be difficult and confusing. He stated that *when the words are not clear he attempts to understand them from the context*.

5.2.9.5 Affective Responses

The following Table (25) illustrates the emotions that emerged from the think aloud protocol of Test-taker I as well as the indicators of these emotions.

Table (25): Affective responses of Test-taker I

Emotion	Indicators from the Think Aloud Data
Disapproval	He expressed disapproval about the writing task stating that a writing on a social topic rather than a letter would be enable him to write more.
Anxiety	He experienced some anxiety in the listening and reading sections: Listening: In Part C he stated that he did not understand several times and at the end he stated that he found it confusing. Reading: He started the section by stating that it looked difficult. He noted several times that it was difficult.
Confusion	He stated that he was confused in Part C of the listening section (he did well), in some items in the grammar section (he did well) and in the last reading passage.
Tiredness	He yawned a lot in the reading section.

5.2.9.6 Fit of the LTP Model

The description of the fit of the LTP model is similar to that of the previous test-takers' profiles with no new patterns emerging from the data. The test-taker only reported metacognitive strategy use in the writing and reading sections.

5.2.10 Test-taker J (Female, advanced English language proficiency stage, 18-22 age range)

5.2.10.1 Writing

The test-taker read the prompt in Arabic and asked, "What is the point of the English version?" She stated that the Arabic prompt was very helpful. The test-taker was pleased that she could copy *some words from the prompt*. She debated whether to write the quantities mentioned in the prompt in numbers of letters. She decided to write once in numbers and one in letters. *She stated her ideas in Arabic and wrote in English*. The test-taker noted that a test-taker can write more on a general topic. She considered the task of writing a letter to be good because she could copy from the prompt without providing new ideas. She thought that a writing task on a general topic will test writing better because test-takers will write in their own English. The test-taker wrote a paragraph that consisted of run-on sentences with grammatical mistakes. There were unsuccessful attempts at compound and complex sentences and some communication of ideas.

5.2.10.2 Listening

In Part A the test-taker did well in this part and got 11 items (out of 15) correct. She was initially surprised that listening was included in the test and stated that "we are not used to listening." She commented that the *directions in Arabic were helpful*. In spite of getting particular items correct the test-taker would note that it was difficult or confusing. At the end of this part, she stated that the listening was difficult even for advanced levels and that the utterances should be repeated twice. In Part B she had more difficulties and got only 9 items correct. She found some items to be difficult or confusing and noted several times that the items should be repeated twice. In parts A & B she noted several times that particular items would be difficult for novice / beginner

level students or even advanced students. In Part C the test-taker got 7 items correct (out of 10) and by that point her anxiety was very high. She started this section by stating that “It is boring. It is boring. I think MCQ is easier. An interview would be better instead of listening. Some people might be good at reading or listening.” She thought that this format of testing listening doesn’t accurately assess the level of students. The test-taker stated several times that the items should be repeated twice. She noted that some items were easy and some were difficult in this section.

5.2.10.3 Grammar

The test-taker got 13 items correct (out of 15) in Part A and 6 items (out of 15) correct in Part B. She believed that although grammar needs concentration, it is easier than the listening section. She stated that all students like MCQ, however, many people get confused and therefore test-takers have to concentrate and read the options carefully because some of the options are confusing. She found Part B to be very difficult even for advanced students. She stated that there should be fewer items of this type because it is confusing. She stated that “a lot of students will find it too difficult, even me” and that a lot of students might leave this part unanswered because it is very confusing. She repeated the word confusing many times. She did not mention the strategies she used.

5.2.10.4 Reading

The test-taker had a negative attitude and considered reading not to be easy. She complained and considered the reading to be unfamiliar and confusing. She also noted that it is not easy for beginners and that there are too many reading passages. The main strategy she used was *rereading*. She got 10 items (out of 20) correct in this section.

5.2.10.5 Affective Responses

The following Table (26) illustrates the emotions that emerged from the think aloud protocol of Test-taker J as well as the indicators of these emotions.

Table (26): Affective responses of Test-taker J

Emotion	Indicators from the Think Aloud Data
Weak Approval	She considered the writing task to be OK but thought that writing on a general topic rather than a letter would assess students' writing ability better.
Surprise	She expressed surprise that a listening task was included in the test.
Anxiety	She experienced anxiety in the listening, grammar (Part B) and reading sections: Listening: She stated many times throughout this section that the utterances should be repeated twice. She stated several times that items were difficult or confusing or that they are difficult for many students. Grammar: She stated that Part B was very difficult and confusing. Reading: She stated that the reading was not easy and confusing. She commented that there too many passages. She handled anxiety by: persisting and rereading.
Boredom	She expressed boredom in Part C on the listening task.
Frustration	She expressed frustration in the listening section stating that an interview would be better and that MCQ items do not accurately test her level.
Confusion	She stated that she felt confused in all sections of the test.

5.2.10.6 Fit of the LTP Model

The description of the fit of the LTP model is similar to that of the previous test-takers' profiles with no new patterns emerging from the data.

5.2.11 **Test-taker K** (Male, advanced English language proficiency stage, 23+ age range)

5.2.11.1 Writing

The test-taker *read the prompt in English and translated as he read*. He then *summarized the writing task in Arabic*. He *formulated his ideas in Arabic and then translated them into English*. He noted that he should write in the past tense. Regarding *word choice, he would pose several options before deciding on the final word to use*. He used the same strategy for sentences. The test-taker wrote fluently with some grammatical mistakes. There were generally successful attempts at complex sentences with some run-ons. There was good communication of ideas.

5.2.11.2 Listening

In Part A the test-taker was very confident and *answered the questions directly*. He did well and got 14 items (out of 15) correct. He was so confident that he found the pace too slow at times. He also did well in Part B and got 14 items (out of 15) correct. He used several strategies such as *reading ahead the options of the following item before listening to the question, translating options and repeating what he heard*. In Part C he got 9 items correct (out of 10) and used the same strategies as in Part B. In item 40 he *read the options and excluded each option one by one*.

5.2.11.3 Grammar

The test-taker was confident. The test-taker got 14 items correct (out of 15) in Part A and 10 items correct (out of 15) in Part B. In Part A the main strategy used was to *read the item and translate*. He used background knowledge: lexical (example: in item 50 the answer is 'shouldn't', since "he is nervous then don't talk to him"). In Part B he would read the item and then exclude options. The test-taker left five *items* in both parts, (items 52,53,58,60 & 63) *to review later* and he reviewed them after completing item 66. He *monitored his time* and commented that he was going at a good pace.

5.2.11.4 Reading

He *read the directions*. He *read the passage and translated parts of it into Arabic*. In the last passage he *summarized the ideas of the passage as he read*. He left 3 *items to be reviewed at the end*. The strategies he used to answer the questions were: *reading the item and translating, reading the item and options, translating the options, excluding options, formulating an answer in Arabic and then selecting an option, excluding options, and initially selecting 2 options and then choosing one option as the correct answer*. He *monitored time* and praised himself (told himself well done) when he found he was in good time.

5.2.11.5 Affective Responses

The following Table (27) illustrates the emotions that emerged from the think aloud protocol of Test-taker K as well as the indicators of these emotions.

Table (27): Affective responses of Test-taker K

Emotion	Indicators from the Think Aloud Data
Confidence	The test-taker was confident in all sections of the test: Listening: He complained several times that the pace was too slow. Grammar: He had sufficient confidence to leave items he was not sure of to review later at the end. Reading: In the last passage he noted that he had a lot of time left and he praised himself.
Self-satisfaction	He praised himself in the last passage in the reading section.

5.2.11.6 Fit of the LTP Model

The description of the fit of the LTP model is similar to that of the previous test-takers' profiles with no new patterns emerging from the data. Test-taker K did not report any metacognitive strategies used. It is interesting to note that this test-taker did not report any negative affective responses while taking the test. Emotional regulation process used was monitoring the time (task-focusing). It is difficult here to interpret the one statement of self-encouragement or self-praise by the test-taker as an indication of emotional regulation of positive emotions. It can also be interpreted as an indication of self-satisfaction as shown in the table above.

5.2.12 Test-taker L (Female, advanced English language proficiency stage, 23+ age range)

5.2.12.1 Writing

The test-taker read the prompt in English twice and then summarized the writing task in Arabic. She wrote a memo instead of a letter since it was to be sent to her boss. She wrote the headings: To, From, Subject Date and then stopped. She stated “how can I start? Let’s think of an appropriate introduction.” She then started writing. She *formulated her ideas in English* and wrote them down. She *reread what she wrote and changed some phrases*. The test-taker wrote fluently with few grammatical mistakes. There were generally successful attempts at complex sentences with good communication of ideas.

5.2.12.2 Listening

In Part A the test-taker did well and got 14 items (out of 15) correct. Although she commented that some pictures did not seem to be clear, however it did not affect her performance. In Part B she had no difficulties in this part and got all 15 items correct. In item 27 she did not pay attention and she had to *guess* (her answer was correct). In Part C she got all 10 items correct.

5.2.12.3 Grammar

The test-taker got 14 items correct (out of 15) in Part A and 12 items (out of 15) correct in Part B. In Part B she *reread some items and repeated some part of a sentence several times*. She also *guessed* some answers.

5.2.12.4. Reading

The test-taker got 16 items (out of 20) correct in this section. She *read the passage and translated portions of it while she was reading*. When answering the questions she *read the items and options and reread them* when required. Any calculations required or numbers mentioned were done in Arabic. In one item she *initially selected 2 options and then chose one option as the correct answer*.

5.2.12.5 Affective Responses

The following Table (28) illustrates the emotions that emerged from the think aloud protocol of Test-taker L as well as the indicators of these emotions.

Table (28): Affective responses of Test-taker L

Emotion	Indicators from the Think Aloud Data
Confidence	The test-taker was confident in all sections of the test: Listening: In question 39 the answer was related to using computers and the test-taker commented that this is her strength. In the grammar and reading sections the test-taker focused on the test items and would justify her answer in some questions.

5.2.12.6 Fit of the LTP Model

The description of the fit of the LTP model is similar to that of the previous test-takers' profiles with no new patterns emerging from the data. Similar to Test-taker K, Test-taker did not report any metacognitive strategies used. It is also interesting to note that both test-takers did not report any negative affective responses while taking the test. In the listening task, Test-taker L's comments about some of the items such as: "Macdonalds makes fries and fish like the one in the picture" and "this is my strength" (she made this comment where the answer to the question was: can use the computer well, meaning that she has good computer skills) reflected her concentration and focus on the task.

5.3 **SUMMARY OF FINDINGS FROM THE THINK ALOUD DATA**

5.3.1 **Test-taking Strategies**

Based on the detailed analysis of the think aloud verbal report data, Table (29) shows all the different cognitive strategies used by the twelve test-takers when responding to a writing task and when taking a multiple-choice test. The strategies were grouped according to the skill being tested: strategies used while writing, general strategies for responding to multiple-choice test items and strategies used specifically when responding to listening and reading comprehension multiple-choice items. Test-takers A, B, C, & D are at the elementary level, Test-takers E, F, G, & H are at the intermediate level while Test-takers I, J, K, & L are at the advanced level.

Table (29): Test-taking strategies identified from the think aloud data

TEST-TAKING STRATEGIES	TEST-TAKERS											
	A	B	C	D	E	F	G	H	I	J	K	L
Specific Strategies for Writing												
Reading the prompt in English only									√		√	√
Reading the prompt first in Arabic and then in English	√											
Reading the prompt first in English and then in Arabic		√	√	√	√	√		√				
Reading the prompt in Arabic only										√		
Rereading the prompt						√						√

Table (29): Test-taking strategies identified from the think aloud data (continued)

TEST-TAKING STRATEGIES	TEST-TAKERS											
	A	B	C	D	E	F	G	H	I	J	K	L
Formulating ideas in Arabic and then translating them into English to write them down	√	√		√	√				√	√	√	
Using words from the prompt	√	√		√	√	√				√		
Formulating ideas in English				√				√			√	√
Rewriting words or phrases											√	√
General MCQ Strategies												
Translating the question & / options	√	√	√	√		√		√	√		√	√
Reading the questions & options	√	√	√	√	√	√	√	√	√		√	√
Selecting an answer & then changing response	√											√
Guessing	√	√	√		√	√	√	√	√			√
Reading instructions		√	√	√	√			√			√	
Leaving answers blank	√	√	√	√	√							
Using background knowledge		√				√	√				√	
Inserting each option one at a time in the question		√			√	√						
Giving up	√	√	√		√	√		√				
Eliminating options				√							√	
Monitoring time					√		√				√	
Running out of time		√		√								
Rereading the questions and options							√	√		√	√	√
Getting stuck on a question							√					
Going back and reviewing answers			√								√	
Producing their own answer before looking at the options provided											√	
Specific Strategies for Listening												
Using knowledge of grammar to answer the question							√					
Reading the options of the next question						√					√	
Specific Strategies for Reading												
Reading the passage first	√	√	√	√	√	√		√	√	√	√	√
Reading the questions first						√	√					
Guessing meaning of an unknown word from the context	√	√					√					
Pronouncing or sounding out words to find their meaning	√											
Rereading parts of the passage			√					√	√			√
Rereading the whole passage								√		√		

Table (29): Test-taking strategies identified from the think aloud data (continued)

TEST-TAKING STRATEGIES	TEST-TAKERS											
	A	B	C	D	E	F	G	H	I	J	K	L
Translating relevant parts of the passage for understanding		√		√					√		√	√
Looking for portion of the passage that the question refers to					√							
Reading the passage and then summarizing the ideas in Arabic											√	

On examining Table (29) it can be seen that the test-takers used a variety of general and skill specific strategies with both similarities and differences in strategy use between high ability (advanced) and low ability (elementary) test-takers. Because the test-taking strategies identified were validated in the second stage of data analysis using a detailed taxonomy to code the data once again (see section 3.10.2.2 in chapter 3) the data in Table (29) was not analyzed in depth. A detailed analysis of the test-taking strategies identified is found in section 5.5.

The think aloud data analysis also showed that it was difficult to classify some of the test-taking strategies used as either contributory or non-contributory. For example, the two strategies: translation and guessing were contributory in some items and non-contributory in other items across the three proficiency levels. Therefore, the LTP model has to be changed to reflect this and the model will no longer differentiate between contributory and non-contributory strategies. The revised LTP model is shown in Figure (7) shown on page 261. It is also interesting to note that there was both agreement and disagreement with the literature regarding contributory and non-contributory strategies. For example, Nevo (1989) identified “selecting an alternative by eliminating others” as contributory and for Test-taker D this strategy was also contributory. Looking for clues in the text was also identified as contributory in Nevo’s study, however for Test-taker E it was both contributory and non-contributory.

5.3.2 Emotions and Emotional Regulation Processes

5.3.2.1 Assessment: metacognitive strategies

The following is a summary of the metacognitive strategies obtained from the VPA data:

- negatively assessing language ability in a specific skill or task;
- positively assessing language ability in a specific skill or task;
- negatively assessing language ability in a specific skill attributed to educational background;
- negatively assessing a specific feature in the design of the test task;
- positively assessing a specific feature in the design of the test task;
- predicting that the task would be easy because of familiarity with the test task format;
- predicting that the task would be difficult because of length (too long);
- predicting that the task would be difficult because of test task format;
- predicting that items would be difficult towards the end of a section;
- comparing current test situation with a previous test-taking experience;
- comparing the current test design to another test;
- deciding that a test task was beyond the test-taker's level of language ability.

..

The metacognitive strategies used consist of test-takers' perceptions of their language ability in specific skills, recalling their previous test-taking experiences, analyzing the design of the test by comparing to other tests they have taken, analyzing features in the test tasks based on their perceptions and predicting the difficulty or ease of a test task in light of their previous test-taking experiences. It is difficult to compare the different metacognitive strategies used among the twelve test-takers because two test-takers did not report any metacognitive strategy use and another two test-takers only provided limited reports of metacognitive strategy use.

Based on the think aloud data, metacognitive strategy use is a set of assessment processes focusing on: self-assessment of language ability and assessment of the test design and test task features. Thus, it depends on test-takers' perceptions of their own

ability and on their previous test-taking experiences. The LTP model postulates that the four sources of information that contribute to the assessment process are: language knowledge, topical knowledge (an individual's knowledge of the world), personal characteristics (age, gender, nationality, resident status, native language, educational background and previous experience with tests) and goals. The findings support the LTP model indicating that metacognitive strategy use does depend on these four sources. Test-takers' assessment of their language ability does depend on both their language and topical knowledge. Moritz (1995) found that when evaluating their own language ability learners use a variety of benchmarks which are related to social category, meaningful others and autobiographical and social contexts. Metacognitive strategy use is dependent on personal characteristics as shown by the think aloud data where education background and previous test-taking experiences emerged as key sources in the assessment process. While the think aloud data did not yield any direct information regarding test-takers' goals, however, in order for an individual to agree to participate in this study and take a placement test, it must be assumed that the goal for doing so is relevant. Furthermore, a goal orientation must be inferred because the think aloud data showed that all test-takers experienced either positive and / or negative emotions. For emotions to emerge, the test must be judged as an important and relevant goal. Therefore, the data does support the LTP model's representation of the assessment component (metacognitive strategy use) based on the four sources of information.

5.3.2.2 Emotions

As a result of the appraisal / assessment process, emotions emerge in a test-taking situation. Table (30) shows the range of emotions or affective responses expressed by the twelve test-takers while they were taking a test.

Table (30): Emotions identified from the think aloud data

EMOTIONS	TEST-TAKERS											
	A	B	C	D	E	F	G	H	I	J	K	L
Positive affective responses												
Pleasure			√									
Confidence	√										√	√
Surprise	√									√		
Approval	√	√	√	√	√	√	√	√		√		
Relief		√					√					
Self-satisfaction											√	
Negative affective responses												
Anxiety	√	√	√	√	√	√	√	√	√	√		
Anger	√	√	√	√		√						
Confusion	√	√		√	√	√	√	√	√	√		
Disappointment	√											
Disapproval		√			√	√	√	√	√			
Boredom			√			√				√		
Complaining					√		√					
Wishful thinking							√					
Helplessness								√				
Tiredness					√		√		√			
Frustration	√	√		√	√	√	√	√		√		

It can be seen from Table (30) that Test-takers K & L only reported positive emotions while Test-taker I only reported negative emotions. All other test-takers reported a considerable range of both positive and negative affective experiences. Pekrun et al (2002) conducted an extensive literature review to identify the range of achievement-related emotions documented. Having identified a lack of empirical research in this area, they listed the following emotions from the literature search they conducted: joy, enthusiasm, hope, relief, pride, gratitude, admiration, sadness, anger, anxiety, hopelessness, shame and guilt, disappointment, envy, boredom, envy, contempt, surprise. It is interesting to note that fear was not included in the list. Nine of the eighteen emotions reported in the literature were identified from the think aloud data. The emotions that were not reported by the test-takers were: joy, enthusiasm, hope, gratitude, admiration, sadness, shame and guilt, envy and contempt. It should be noted that the emotions identified in the literature were based on research conducted in the area of learning and achievement and not necessarily in test-taking conditions.

Therefore, it is expected that the affective responses obtained from the think aloud data would be somewhat different from emotions related to instruction and the process of studying. It is understandable in a test-taking situation that test-takers would not be likely to report joy and enthusiasm, however, test-takers did report that they were pleased with certain aspects of the test. Gratitude and admiration can be related to "approval" that emerged from the think aloud data. Hope was not directly expressed by the test-takers. They predicted that a certain section of the test would be easy, however, they did not explicitly mention hope. Shame and guilt were not mentioned at all. Shame is a strong emotion and is not likely to occur within a test-taking situation. Furthermore, it is an emotion associated with 'others' so it is unlikely to be an issue in the context of an actual test. It may emerge later on when the test results are announced. Guilt may occur when a test-taker feels that he / she is not well prepared. Envy and contempt are also very strong emotions and it is not likely that Egyptian learners would express these feelings. It is not culturally appropriate to do so. The emotions that emerged from the think aloud data that were not mentioned by Pekrun et al (2002) were: approval, confusion, disapproval, complaining, wishful thinking, tiredness and frustration.

Schutz et al (2002:336) state that "current research on appraisal indicates that feeling confident, feeling in control, and feeling certain will result in potentially feeling enjoyment, pride or satisfaction" and this was clearly shown in Test-taker K's think aloud. He was clearly in control and he was satisfied with his performance by praising himself.

Having identified the affective responses of each test-taker, the researcher analyzed the think aloud data to infer the processes used to handle these emotions. The LTP model postulates that when positive emotions emerge, the test-taker uses task-focusing emotional regulation processes such as monitoring time, self-encouragement and keeping negative thoughts away. When negative emotions emerge, the coping processes used by a test-taker would be self-blame, self-criticism, self-talk and taking deep breaths. These emotion-focusing processes help test-takers cope with the stressful situation or negative emotions.

5.4 THE APPLICABILITY OF THE LTP MODEL

In light of the think aloud data analysis as discussed, the LTP is applicable to a large extent and describes the processes involved in a test-taking situation. However specific components of the model have to be revised to explain the data obtained. The model will be reviewed once again in light of the results from the quantitative data analysis (the final model is discussed in section 7.2.4). The think aloud data showed that when there was a positive affective response, it was difficult to identify any processes used to regulate the positive emotions. The data showed that when the test-taker experienced positive emotions, he / she would continue what they are doing (applying test-taking strategies) thus, stabilizing a particular behavioral pattern or would focus on specific aspects of the task at hand such as commenting that the listening test was very good or saying nothing at all (task-focusing processes). This was clearly seen in Test-takers' K and L think aloud protocols where all the data focused on cognitive strategies use and thoughts related to the different tasks and items.

The think aloud data also showed that it was difficult to separate emotion-focusing processes from emotions. The results showed that when a negative affective response occurred there were two dimensions: the test-taker experienced several negative emotions and applied task-focusing processes. The test-taker also used test-taking strategies. Thus, emotional regulation is the interaction between assessment, emotions and task-focusing processes. The task-focusing processes identified were: focusing on the format of a letter, focusing on the presence of a prompt in Arabic and English, focusing on positive aspects of the test, focusing on the instructions in Arabic and managing time.

From the data it was not possible to classify test-taking strategies as contributory or non-contributory. Therefore, the changes to the model are: the test-taking strategies are not classified as contributory and non-contributory; in case of both a positive affective response, task-focusing processes replace emotional regulation processes; and when a negative affective response occurs the emotional regulation processes (replaces coping processes) consist of the interaction between emotions and task-focusing processes.

The LTP model in light of the think aloud data is described as follows.

In a test-taking situation, the test-taker evaluates his / her personal capabilities to complete a particular test task. The test-taker analyzes the requirements of the test task. In assessing ability to perform this task, the test-taker evaluates his / her language ability in relation to his / her topical knowledge and personal characteristics. The results of these assessments may or may not lead to a discrepancy between the requirements of the test task and language knowledge, topical knowledge and personal characteristics. Parallel to this process, the test-taker also makes comparisons between his / her goals and the test situation. If there is no discrepancy and the test is judged as not important to the test-taker's goals, there is no positive or negative affective response. If there is a discrepancy and the test is judged as being an important and relevant goal, it may lead to a positive or negative affective response. As a result positive or negative emotions occur where test-takers apply test-taking strategies and task-focusing processes which ultimately lead to successful or unsuccessful test task performance. When there is a positive affective response (the mastery mode), positive emotions are experienced and the test-taker uses test-taking strategies and task-focusing processes on the test tasks. In case of a negative affective response (the coping mode), negative emotions emerge and task-focusing processes are used. In both modes the test-taker selects a variety of test-taking strategies to be used (there is no differentiation between contributory and non-contributory strategies). The interplay between the test-taking strategies selected, the emotions that emerge and the task-focusing processes used determines the behavior of the test-taker. Test-takers may increase or decrease effort, persist in completing the task, give up or seek help (ask for answers or even attempt to cheat during the test). The consequences of test performance are either successful or unsuccessful experiences which feedback once again into the sources of information used for assessing capability.

The test-taker's behavior during the test-taking process is the outcome of an initial assessment followed by the interplay between the selection and adaptation of test-taking strategies, the emotions that emerge and of task-focusing processes. Emotional regulation consists of assessment and task-focusing processes and emotions which interact with test-taking strategies. This interaction between test-taking strategies and emotional regulation related to test performance will be explored further in Research Question 3.

5.5 FINDINGS FROM THE QUANTITATIVE ANALYSIS OF THE THINK ALOUD DATA

5.5.1 Frequency of Test-taking Strategies Used as Reported by Test-takers

The think aloud data were coded and the results are shown in Table (31). The codes were generated from both the literature and the think aloud data. Table (31) shows the frequency of occurrence of the strategies reported by each test-taker. The frequency counts are those that were assigned by the two raters separately or that were reached by consensus. When analyzing the frequency counts it is important to note that these numbers cannot used to make statistical comparison since some of the strategies are not countable: they are either used or not. For example a test-taker would “read the prompt in English only” or not. In some of the think alouds the test-takers mentioned that they used a particular strategy for several items but they were not specific. For example, a test-taker stated that she had to guess the answer for several items and she did not mention these specific items. Here the strategy used: “guessing” is coded as one occurrence. The frequency counts will be treated as indicators of patterns underlying the data. The last two columns show the total number times each strategy was reported (ST) by the number of test-takers (TT) who did so. The last row shows the number of different types of strategies reported by each test-taker (not the total number of strategies used by each test-taker since this may be misleading). The description of the codes is found in section 3.10.2.2 of Chapter 3.

Table (31): Frequency count of test-taking strategies reported

TEST-TAKING STRATEGIES	TEST-TAKERS												TOTALS	
	A	B	C	D	E	F	G	H	I	J	K	L	ST	TT
TTS1	19	3	22	23	9	8	11	24	30		44	52	245	11
TTS2	2		1										3	2
TTS3				1							10		11	2
TTS4	2	12	3	31		4		4	9		24	1	90	9
TTS5		1			2	1							4	3
TTS6		1				1	1				1		4	4
TTS7											1		1	1
TTS8							1	6		1	2	7	17	5
TTS9											7		7	1
TTS10	18	18	5	22	1								64	5

Table (31): Frequency count of test-taking strategies reported (continued)

TEST-TAKING STRATEGIES	TEST-TAKERS												TOTALS	
	A	B	C	D	E	F	G	H	I	J	K	L	ST	TT
TTS11	5	4	1		1	7	4	2	1			2	27	9
TTS12	1											1	2	2
TTS13														
TTS14														
TTS15		1		1									2	2
TTS16														
TTS17														
TTS18					1		1				1		3	3
TTS19														
TTS20		2	2	1	2			1			1		9	6
TTS21									1	1	1	1	4	4
TTS22			1								1		2	2
TTS23							1						1	1
TTS24									1		1	1	3	3
TTS25	1												1	1
TTS26		1	1	1	1	1		1					6	6
TTS27										1			1	1
TTS28						1						1	2	2
TTS29	1	7		1	1				1	1	1		13	7
TTS30	1	1		1	1	1				1			6	6
TTS31				1				1			1	1	4	4
TTS32											1	1	2	2
TTS33						1					7		8	2
TTS34							1						1	1
TTS35	4	2	1	1	1	1		3	1	1	5	1	21	11
TTS36						2	1						3	2
TTS37								1		1			2	2
TTS38			1					1	1			1	4	4
TTS39		4		3				1			4	1	13	5
TTS40											1		1	1
TTS41	1	1					1						3	3
TTS42														
TTS43	1												1	1
TTS44														
TTS45					1								1	1
TTS46														
TTS47														
TTS48														
TTS49														
Total types of strategies	12	14	10	12	11	11	9	11	8	7	19	13		

On comparing the data obtained in Table (29) and Table (31) it can be seen that there are some differences. Strategies TTS2, TTS9, TTS13, TTS14, TTS16, TTS17, TTS19, TTS42, TT44, TTS46, TTS47, TTS48 and TTS49 were not identified in the first stage of analysis. This can be explained that the data in each table was obtained by a different approach. An inductive approach was used to identify the strategies in Table (29) while a deductive approach was used in Table (31). Thus, the outcome of the inductive analysis was that some strategies were missed either because they were not used by the test-takers or they were not considered by the researcher. Similarly, there are some differences between Table (29) and Table (31) focusing on emotions identified.

On examining the range of strategies used by the test-takers in Table (31), it can be seen that Test-taker K at the advanced level used the most variety of strategies while Test-taker J also at the advanced level used the least range of strategies. Furthermore, Test-taker B at the elementary level used fourteen different strategies. Thus, the range or number of different strategies used does not seem to be related to proficiency level.

From both the total frequency counts and number of test-takers reporting using a particular strategy, the four most commonly used general strategies (TTS1 – TTS23 are the general strategies on the TTSQ) across the three proficiency levels were reading the questions and options (reported by eleven test-takers), translating the question and / options (nine test-takers), guessing (nine test-takers), and reading instructions (six test-takers). Five of the test-takers at the elementary and intermediate levels skipped items by leaving them blank (none of the test takes at the advanced level did so) and only the four advanced level test-takers answered all items and never left an answer blank. Five test-takers at the intermediate and advanced levels reread the question and options for clarification. Four test-takers at different levels reported using background knowledge. Three test-takers (intermediate and advanced levels) monitored time and three test-takers (at the elementary and intermediate levels) reported inserting each option at a time in the question. Only one or two test-takers reported using each of the following strategies: stopping reading options, producing their own answer before reading the options, postponing dealing with a question, changing responses, running out of time, reviewing

answers and getting stuck on one question. Thus, there are some differences in strategy use between elementary and advanced level test-takers.

In the writing task (TTS24 – TTS32) the most commonly used strategy across the three proficiency levels (seven test-takers) was formulating ideas in Arabic and then translating them into English to write them down (TTS29). Six of the twelve test-takers at the elementary and intermediate levels read the prompt in English and then in Arabic (TTS26). As expected it was only three of the four advanced level test-takers who read the prompt in English only (TTS24). Unexpectedly, the only test-taker who read the prompt in Arabic only was at the advanced level. The think aloud data showed that Test-taker J experienced high anxiety throughout the test and she performed poorly in all the tasks. Test-taker J was also the only advanced level test-taker who copied words from the prompt (TTS30) while five of the eight test-takers at the elementary and intermediate levels reported using this strategy. Four test-takers across the three proficiency levels reported formulating their ideas in English and two advanced level test-takers reported rewriting words or phrases. One test-taker at the elementary level read the prompt first in Arabic and then in English and two test-takers reported rereading the prompt.

Regarding specific strategies related to listening, two test-takers reported reading the options of the following question and one test-taker at the intermediate level reported using grammar when responding.

In the reading task, almost all test-takers (except for one test-taker) read the passage and then the questions (TTS35). Two test-takers read the questions before reading the passage (TTS36). Five test-takers across the three proficiency levels translated parts of the passage while they were reading (TTS39). Four test-takers reported rereading parts of the passage (TTS38) and two test-takers reported rereading the whole passage (TTS37). It is important to note that none of the elementary test-takers were able to complete the reading section and gave up without rereading. Three test-takers reported using the context to identify meanings of words (TTS41). One test-taker at the elementary level pronounced words in an attempt to find their meaning (TTS43), another test-taker at the intermediate level reported searching for those parts of

the passage to which the questions refer (TTS45) and an advanced test-taker summarized the ideas of the passage in Arabic (TTS40).

The twelve test-takers did not report using twelve of the strategies: TTS13, TTS14, TTS16, TTS17, TTS19, TTS42, TTS44, TTS46, TTS47, TTS48, and TTS49. While some of the test-takers may have used some of these strategies, but may have not occurred to them to report their use. Therefore, these strategies must be further investigated by directly asking test-takers whether they use them or not.

While the think aloud data showed that there were some differences in strategy use between high ability (advanced) and low ability (elementary) groups, however for the most part, similar strategies were used across the three ability groups. Variability of strategy use across different ability groups is further investigated in Research Question 3 described in Chapter 6.

5.5.2 Frequency of Emotions, Emotional Regulation Processes and Behaviors as Reported by Test-takers

Table (32) shows the frequency of occurrence of emotional regulation processes and emotions reported by each test-taker. As previously discussed, the frequency counts will be treated as indicators of patterns underlying the data. The last two columns show the total number of times each process / emotion / behavior was reported (EM) by the number of test-takers (TT) who did so. The range of emotions identified was discussed in section 5.3.2.2 above. The description of the codes is found in section 3.10.2.2 of Chapter 3.

Table (32): Frequency count of emotional regulation processes / emotions / behaviors reported

EMOTION/ PROCESS*	TEST-TAKERS												TOTALS	
	A	B	C	D	E	F	G	H	I	J	K	L	EM	TT
AFF1	1	1	1										3	3
AFF2								2					2	1
AFF3	10	5	3	4	2	2	3	2		2	3	2	38	11
AFF4	8	27	8	9	3	9	12	15	8	23			122	10
AFF5	1	1		3	6	5	3	3	3	20			45	9
AFF6							1						1	1
AFF7														
AFF8		2						1					3	2

Table (32): Frequency count of emotional regulation processes / emotions / behaviors reported (continued)

EMOTION/ PROCESS*	TEST-TAKERS												TOTALS	
	A	B	C	D	E	F	G	H	I	J	K	L	EM	TT
AFF9														
AFF10	1		3		1	2	2	1	2		5		17	8
AFF11														
AFF12	2	2	3		1		3	2		2			15	7
AFF14		1					2						3	2
AFF15			2	2		1	5	2					12	5
AFF16	2	1			2		1			1			7	5
AFF18								1					1	1
AFF19					1		1		2				4	3
AFF20	2	1	1	1		1		2					8	6
AFF21					1		1						2	2
AFF22									1				1	1
AFF23		3	1	2		7	2	2			2	9	28	8
AFF24														
AFF25	2		1			1				1			5	4
AFF26														
AFF27	1	5		2	1	4	1	1		1			16	8
AFF28	1		2	1		1		1		1			7	6
AFF29						2							2	1
AFF30		3											3	1
AFF31	1	1	1		1	3		1		1			9	7

* The numbering of the codes is not correct (after having coded the data, 13 & 17 were discovered to be missing).

As shown in Table (32), the three most frequently occurring emotional regulation processes and emotions are: feeling that the test or item was difficult (AFF4), feeling that the test or item was confusing (AFF5) and feeling that the test or item was easy (AFF3). Two of these are appraisal or assessment processes and one is an emotion. The next most frequently occurring processes reported by eight test-takers were not being sure of the correct answer (AFF23: task-focusing), having enough time (AFF10: task-focusing) and feeling frustrated due to lack of time (AFF27: emotion). It should be noted that Test-takers A, E, F, G and H reported that there was sufficient time in one section of the test and not enough time in other sections. Test-takers A, E, F and H reported that there was sufficient time for the writing task but not enough time in one or more of the other three sections of the test. Test-taker G initially reported that in part A

of the listening section, the timing was adequate but in the second and third parts of the task as anxiety increased, he felt frustrated due to lack of time.

The most frequently occurring emotions / processes were distributed across the three proficiency levels. The only two emotions / behavior that showed some variation in use related to proficiency level were feeling nervous during the test (AFF15: emotion) and giving up because the test was too difficult (AFF20: behavior). This emotion and behavior were reported only by elementary and intermediate test-takers and not the advanced group.

Seven test-takers considered the test to be a challenge (emotion) and another seven felt that it was a good idea to have the prompt in both Arabic and English (task-focusing). Six test-takers also felt that it was a good idea to have the instructions in both Arabic and English (task-focusing) and five test-takers had difficulty in concentrating (emotion). The rest of the processes were reported by four test-takers or less.

The twelve test-takers did not report using any of the following five processes: AFF7, AFF9, AFF11, AFF24, and AFF26. Three of these emotions / processes are related to enjoying the test experience and having a goal orientation. As previously mentioned, some of the test-takers may have used some of these processes or experience these emotions, but did not think of reporting their use. Thus, these emotions / processes / behaviors must be further investigated by directly asking test-takers whether they use them or not.

Therefore, the most frequently occurring affective response are appraisal processes and the feeling of confusion. Based on this particular data set there is very low variation in emotional regulation process / emotions across the three proficiency levels.

As described in the Research Design chapter the methodology adopted in this study is a mixed methods approach and thus, in order to address Research Question 2 quantitatively a Likert-type questionnaire was administered to test-takers after they completed the English language placement test. The results of this phase are reported in the following sections. Four hundred and ninety seven (497) respondents completed the Test-taking Strategies Questionnaire (TTSQ). The TTSQ consisted of 49 closed items focusing on test-taking strategies and 29 closed items on emotions that occurred during

the test. The TTSQ is described in detail in Chapter 3. The results from administering the TTSQ addressing Research Question 2 are described in the following sections.

5.6 FINDINGS FROM THE TTSQ DATA

5.6.1 TEST-TAKING STRATEGIES

5.6.1.1 Test-taking Strategy Use

Table (33) shows the means and standard deviations (SD) for each of the TTSQ items related to test-taking strategies. Test-takers responded using a 3-point rating scale: 3 = I used this strategy several times during the test; 2 = I used this strategy only once or twice during the test; 1 = I did not use this strategy during the test. Thus, a '3' indicates high frequency of use while a '1' indicates low frequency of use. When interpreting the means of these items, a mean of 2.5 or higher indicated a high frequency test-taking strategy. Means ranging from 1.5 to 2.4 indicated low frequency test-taking strategies and means of less than 1.5 indicated that the strategy is not used or hardly used at all. High frequency strategies are indicated in bold and strategies not used or hardly used at all are shaded in the following table.

Table (33): Means & SD of the testing taking strategies in the TTSQ

#	Test-taking Strategy	Mean	SD
1	Read the questions and options before choosing one.	2.6	0.65
2	Stopped reading options when I got to the one that seemed correct.	1.9	0.79
3	Translated the question and the options.	2.1	0.84
4	Selected the option by eliminating the other 3 options.	2.0	0.76
5	Inserted each option one at a time in the question.	1.8	0.73
6	Made an educated guess using background knowledge.	2.1	0.70
7	Tried to produce own answer to the question before looking at the options provided.	1.5	0.71
8	Reread the questions and options for clarification.	2.5	0.65
9	Postponed dealing with a question or selecting a given option until later.	1.4	0.62
10	If there was a question that was not understood, left the answer blank.	1.6	0.82
11	Guessed without any particular considerations.	1.6	0.69
12	Changed answer when appropriate.	2.1	0.62
13	When not sure of the answer, selected an option that was longer / shorter than the others.	1.2	0.48
14	When not sure of the answer, looked for an option that seemed to be different from the others.	1.5	0.64

Table (33): Means & SD of the testing taking strategies in the TTSQ (continued)

#	Test-taking Strategy	Mean	SD
15	Ran out of time without trying all the questions.	1.7	0.77
16	Tried to finish the test as fast as possible.	1.9	0.84
17	Started by reviewing or surveying the whole test.	1.4	0.67
18	Monitored the time.	1.7	0.78
19	Watched to see when other students finish the test.	1.2	0.56
20	Read the instructions carefully.	2.7	0.60
21	Did not leave any answer blank.	1.7	0.80
22	Went back and reviewed or checked my answers.	1.6	0.69
23	Got stuck on one question for a long time.	2.2	0.58
24	Read the prompt in English only.	1.8	0.80
25	Read the prompt first in Arabic and then in English.	1.9	0.83
26	Read the prompt first in English and then in Arabic.	1.8	0.79
27	Read the prompt in Arabic only.	1.3	0.61
28	Reread the prompt.	1.9	0.76
29	Formulated ideas first in Arabic and then translated them into English and wrote them down.	1.8	0.83
30	Used words from the prompt in my answer.	2.0	0.71
31	Formulated my ideas in English only.	2.0	0.84
32	Rewrote words or phrases in my answer.	1.8	0.70
33	Had sufficient time to read the options of the next question before listening to the question on tape.	1.9	0.77
34	Used knowledge of grammar to answer the question.	2.4	0.71
35	Read the passage first.	2.7	0.60
36	Read the questions first before reading the passage.	1.4	0.67
37	Reread the whole passage.	1.7	0.73
38	Reread parts of the passage.	2.3	0.62
39	Translated relevant parts of the passage to understand.	1.9	0.77
40	Read the passage and then summarized the ideas in Arabic.	1.5	0.73
41	Was able to guess the meaning of an unknown word from context.	2.1	0.65
42	Skipped unknown words.	2.3	0.77
43	Pronounced or sounded out the word to find its meaning.	1.9	0.77
44	Used knowledge of grammar to answer the question.	2.3	0.72
45	Looked for the portion of the passage that the question refers to and then looked there for clues to the answer.	2.4	0.68
46	Matched material from the passage with material in the question and in the options to find the answer.	2.3	0.74
47	Selected an option because it appeared to have a word or phrase from the passage in it.	1.6	0.73
48	Selected an option based on understanding the passage read.	2.6	0.60
49	Got clues from answering one question that were helpful in answering another question.	1.5	0.65

Table (33) shows that the means of most of the strategies (38 of the 49 strategies) ranged from 1.5 to 2.4. This means that the majority of respondents reported occasional use of these strategies or a significant number of respondents reported not using this particular strategy. From Table (33) and based on the means, the five (5) most frequently occurring test-taking strategies are:

- reading the question and options before choosing one;
- rereading the question and options for clarification;
- reading the instructions carefully;
- reading the passage first;
- selecting an option based on understanding the passage read.

These high frequency strategies reflect test-takers’ attempts at clearly understanding what is required from the task or question under consideration before responding. In the think aloud data two of these strategies were also identified as high frequency: reading the question and options and reading the instructions carefully. It was surprising that guessing without any particular considerations (TTS11) and translation of questions and options (TTS3) which appeared as high frequency strategies in the think aloud data were reported as low frequency strategies in the quantitative data (the means were 1.6 and 2.1 respectively). This may be that test-takers would be reluctant to report the use of such strategies or that they genuinely believe that they did not use them frequently. On examining the frequency distributions of these two strategies the following results were obtained:

Table (34): Frequency distribution of TTS3 and TTS11 by percentage (N=497)

Strategy	Frequency Distribution on the Rating Scale		
	1	2	3
TTS3	30%	28%	42%
TTS11	48%	40%	12%

Although about one third of the respondents (30%) reported not using translation at all however 42% reported using it several times (high frequency use). About half of the respondents (48%) stated that they did not resort to guessing and 40% reported using it once or twice during the test. Only 12% reported using guessing several times.

Table (33) shows that based on the means, the six (6) test-taking strategies that were reported as not being used or hardly used at all are:

- postponing dealing with a question or selecting a given option until later (TTS9);
- when not sure of the answer, selecting an option that was longer / shorter than the others (TTS13);
- starting by reviewing or surveying the whole test (TTS17);
- watching to see when other students finish the test (TTS19);
- reading the prompt in Arabic only (TTS27);
- reading the questions first before reading the passage (TTS36);

These results were similar to the think aloud data: TTS13, 17 and 19 were not reported, TTS27 was used by one test-taker and TTS9 and TTS36 were used by two test-takers.

5.6.1.2 Qualitative Data on Test-taking Strategy Use

Respondents were asked whether they used a strategy while taking the test that was not mentioned in the questionnaire. Most of the respondents (96%) responded “NO” indicating that they did not use a strategy not mentioned in the questionnaire while 4% (22 respondents) responded “YES” indicating that they did. Those who responded “YES” were requested to mention the strategy and to indicate whether they used it once or twice or more. There were 26 responses to this open ended question and they were analyzed. Only five (5) responses included ‘new’ strategies not mentioned in the questionnaire. The remainder of the responses all included strategies listed in the TTSQ. It is interesting to note here that nine (9) test-takers stated that they relied on guessing. One test-taker at the intermediate level stated “during the listening section I could not hear the sentence so I just guessed the answer. This happened twice to be

honest.” Another test-taker at the elementary level noted that “I used the strategy of guessing for the last questions because I did not have enough time to read the passage well and answer the questions.”

The five new test-taking strategies / behaviors that emerged from the qualitative data were:

- Looking for a word heard on the tape in the options until it is found and then basing the answer on it when the test-taker did not know the answer to the question;
- Stopping to take a break (behavior);
- Recalling information about memorized words;
- Choosing the middle option when the test-taker did not know the answer and did not want to leave the question blank;
- Refusing to answer the question when not finding an answer she was convinced with and not wanting to answer without being sure what it is.

Three of these strategies are coping strategies when the test-taker did not know the answer to the question. In the listening task the test-taker focused on words hoping that one of them would give a clue to the answer. Just focusing on option 2 or 3 is actually guessing and refusing to answer the question means that the test-taker leaves it blank. It is interesting to note that the test-taker who took a rest or a break stated that this resulted in him not being able to complete the test. Memorization is a learning strategy that is predominantly used in the Egyptian education system and thus, it would be expected that test-takers would consider recalling memorized information as a test-taking strategy even on a proficiency test.

5.6.1.3 Results of the Factor Analysis of Test-taking Strategies

In order to reduce the data and identify factors or pattern underlying test-taking strategies, the responses were analyzed using factor analysis. After Varimax rotation, a fifteen (15) factor solution was obtained which accounted for 57.5% of the total variance in the test-taking strategies section of the TTSQ as shown in Table (35).

Table (35): Factor analysis for test-taking strategies

Factor	Label	Eigenvalue	Variance %	Cumulative Variance
Factor 1	Making use of clues	5.8	4.5	4.5
Factor 2	Dealing with unknown texts	3.7	7.4	11.9
Factor 3	Using L1 to deal with L2	2.5	6.2	18.1
Factor 4	Leaving blank responses	1.9	4.7	22.8
Factor 5	Managing time	1.8	2.8	25.6
Factor 6	Using knowledge of grammar rules	1.6	4.1	29.7
Factor 7	Dealing with unknown words	1.5	3.7	33.4
Factor 8	Rereading text	1.4	3.0	36.4
Factor 9	Handling MCQ options	1.3	2.9	39.3
Factor 10	Recognizing correct answers without hesitation	1.2	2.7	42.0
Factor 11	Elimination	1.2	3.0	45.0
Factor 12	Rereading questions	1.1	3.4	48.4
Factor 13	Ensuring understanding	1.1	3.3	51.7
Factor 14	Reading the prompt in two languages: L1/L2/L1	1.1	3.0	54.7
Factor 15	Previewing / reviewing	1.0	2.8	57.5

Tables (36 - 48) describe the questionnaire items that are included in each of the fifteen (15) factors and the item loadings of each. Three items load on Factor 1 as shown in Table (36):

Table (36): Item loadings on Factor 1 (TTS)

TTSQ Item	Loading
Looked for the portion of the passage that the question refers to and then looked there for clues to the answer (TTS45)	0.70
Matched material from the passage with material in the question and in the options to find the answer (TTS46)	0.70
Used words from the prompt in my answer (TTS30)	0.46

Factor 1 has been labeled as “making use of clues” because all three items are related to test-takers’ strategies to use clues to find the answer. These clues are looking at portions of a reading passage, looking at portions of the reading passage that match the question and options and using words in the writing prompt to complete the task: writing a letter.

Factor 2 consists of seven questionnaire items as follows:

Table (37): Item loadings on Factor 2 (TTS)

TTSQ Item	Loading
Read the questions first before reading the passage (TTS36)	0.66
When not sure of the answer, selected an option that was longer / shorter than the others (TTS13)	0.63
Read the passage first (TTS35)	-0.62
Selected an option based on understanding the passage read (TTS48)	-0.53
When not sure of the answer, looked for an option that seemed to be different from the others (TTS14)	0.49
Selected an option because it appeared to have a word or phrase from the passage in it (TTS47)	0.48
Watched to see when other students finish the test (TTS19)	0.46

Initially, this factor was difficult to interpret since it included items from different parts of section 1. Factor 2 was labeled as “dealing with unknown texts” because it was determined that the focus was reading. Four of the items (TTS36, TTS35, TTS48 and TTS47) were all from the strategies for reading section of the questionnaire. The reading comprehension section was the last part of the test and thus, watching to see when other students finish the test is relevant. The inclusion of TTS13 and TTS14 in this factor indicated that the task was difficult and test-takers had to resort to such strategies when not being able to answer. In the context of the Egyptian education system where most of the tests focus on rote memorization, the inclusion of previously unseen reading texts poses a challenge to many test-takers and therefore, “dealing with unknown texts” emerges as a strategy. The negative correlation of the two strategies TTS35 and TTS48 with the others also indicates that it is a strategy used when challenged. Thus, a high score on this factor would indicate that the test-taker found the reading texts to be difficult.

Factor 3 consists of five questionnaire items as follows:

Table (38): Item loadings on Factor 3 (TTS)

TTSQ Item	Loading
Translated the question and the options (TTS3)	0.73
Formulated ideas first in Arabic and then translated them into English and wrote them down (TTS29)	0.68
Translated relevant parts of the passage to understand (TTS39)	0.66
Formulated my ideas in English only (TTS31)	-0.59
Read the passage and then summarized the ideas in Arabic (TTS40)	0.57

This factor is readily interpretable since all the items are related to translation or the use of L1 (Arabic). It is interesting to note the use of English only which is the reverse of resorting to Arabic is negatively correlated with the other four items. Factor 3 was labeled as “using L1 to deal with L2”. Thus, a high score on this factor would indicate high frequency of use of L1.

Three questionnaire items load on Factor 4 as shown in Table (39):

Table (39): Item loadings on Factor 4 (TTS)

TTSQ Item	Loading
Did not leave any answer blank (TTS21)	0.85
If there was a question that was not understood, left the answer blank (TTS10)	0.79
Ran out of time without trying all the questions (TTS15)	0.58

Factor 4 is called “leaving blank responses” because the three items are related to whether or not the test-taker is determined to complete all the answers and not to leave any blanks. This was also illustrated in the qualitative data discussed above regarding new strategies used where the test-taker made a conscious decision to leave blanks when she did not know the answers. Another example taken from the qualitative data obtained from the open-ended question on the questionnaire asking respondents to briefly describe their feelings during the test is: “I felt that it was actually a test that identified the abilities of each person. That is why I answered only what I knew and I left a lot that I did not know.” A test-taker confessed that he guessed the answers to the last questions on the test because he ran out of time.

Factor 5 consisted of only one item and the researcher labeled it as “managing time”. The item is “monitored the time” (TTS18) with a loading of 0.74.

Factor 6 consisted of two items as follows:

Table (40): Item loadings on Factor 6 (TTS)

TTSQ Item	Loading
Used knowledge of grammar to answer the question (TTS34) (listening)	0.81
Used knowledge of grammar to answer the question (TTS44) (reading)	0.77

Factor 6 is very clear and it is labeled as “using knowledge of grammar rules. This strategy is expected to emerge since the grammar-translation approach to English language teaching in Egypt has dominated most EFL classroom methods. In the last few years as a result of efforts to reform the educational system, communicative language teaching methods are being increasingly applied, however, most of the Egyptian adult test-takers who responded to the TTSQ would have most likely experienced the more traditional form of language teaching.

Two items load on Factor 7 as follows:

Table (41): Item loadings on Factor 7 (TTS)

TTSQ Item	Loading
Was able to guess the meaning of an unknown word from context (TTS41)	0.69
Skipped unknown words (TTS42)	0.63

The researcher labeled Factor 7 as “dealing with unknown words”. This focus on words is interesting and could also be a reflection of the test-takers’ educational background. In schools there has been a lot of emphasis on memorizing words and word lists were frequently used by both teachers and students (and probably still are used for high stakes tests such as the school leaving test). This was also illustrated in the qualitative data discussed above regarding new strategies used where a test-taker

specifically stated the use of the strategy of recalling information about memorized words.

Factor 8 consisted of two items as follows:

Table (42): Item loadings on Factor 8 (TTS)

TTSQ Item	Loading
Reread parts of the passage (TTS38)	0.78
Reread the whole passage (TTS37)	0.54

Factor 8 was labeled as “rereading text”. In the think aloud data, several test-takers also reported using this strategy. Once again, repetition is highly encouraged in the educational system and when students prepare for a language test, they read and reread the reading comprehension passages in their assigned textbooks.

Factor 9 also consisted of two items as shown in Table (43):

Table (43): Item loadings on Factor 9 (TTS)

TTSQ Item	Loading
Inserted each option one at a time in the question (TTS5)	0.69
Tried to produce own answer to the question before looking at the options provided (TTS7)	0.62

Factor 9 was labeled as “handling options” because the test-taker would either try out each of the four options one at a time or not look at the options at all and try to produce their own answer. In the think aloud data more test-takers substituted options in the stem compared to initially disregarding the options.

Factor 10 consisted of only one item and the researcher labeled it as “recognizing correct answers without hesitation”. The item is “stopped reading options when I got to the one that seemed correct” (TTS2) with a loading of 0.75. This strategy is probably used by a confident test-taker who is sure of him / herself. It also could be used by an impulsive test-taker who just seizes on to the first answer that seems correct without continuing to check the remaining options to be sure of the correct answer.

Factor 11 consisted of two items as shown in Table (44):

Table (44): Item loadings on Factor 11 (TTS)

TTSQ Item	Loading
Changed answer when appropriate (TTS12)	0.73
Selected the option by eliminating the other 3 options (TTS4)	0.45

Factor 11 was initially not obvious, however, when considering that changing an answer actually means that the test-taker eliminates it in favor of a more correct answer, it became clear. Factor 11 is labeled as “elimination”.

Two items load on Factor 12 as shown in Table (45):

Table (45): Item loadings on Factor 12 (TTS)

TTSQ Item	Loading
Reread the prompt (TTS28)	0.54
Got stuck on one question for a long time (TTS23)	0.50

Factor 12 is labeled as “rereading questions”. Getting stuck on one question for a long time implies that the test-taker rereads the question several times. A high score on this factor indicates that the test-taker tends to reread questions frequently and this may result in running out of time at the end.

Two items load on Factor 13 as shown in Table (46):

Table (46): Item loadings on Factor 13 (TTS)

TTSQ Item	Loading
Read the questions and options before choosing one (TTS1)	0.74
Reread the questions and options for clarification (TTS8)	0.60

Factor 13 is labeled as “ensuring understanding” because a careful test-taker would read the questions and options before choosing one. When understanding is not clear the test-taker would reread the questions and options to ensure understanding.

Factor 14 consists of two items as shown in Table (47):

Table (47): Item loadings on Factor 14 (TTS)

TTSQ Item	Loading
Read the prompt first in English and then in Arabic (TTS26)	0.69
Read the prompt first in Arabic and then in English (TTS25)	-0.54

Factor 14 is related to whether a test-taker reads the writing prompt in English first or in Arabic first. Although the interpretation of this factor is clear it was difficult to find a label that would capture the meaning. It is labeled as “reading of the prompt in two languages: L1/L2/L1. TTS25 is negatively correlated to TTS26 and thus, it can be viewed as a continuum. Similar to the strategy of “using L1 to deal with L2”, a test-taker would most likely use this strategy as support. It was decided that a high score on this factor would indicate that the test-taker would read the prompt first in Arabic.

The two items that load on Factor 15 are shown in Table (48):

Table (48): Item loadings on Factor 15

TTSQ Item	Loading
Had sufficient time to read the options of the next question before listening to the question on tape (TTS33)	0.69
Rewrote words or phrases in my answer (TTS32)	0.45

This was the most challenging factor to interpret because upon the initial reading of the two items there did not seem to be any relationship between them. However, on further analysis an underlying pattern was identified. When a test-taker rewrites words or phrases this means that he / she is reviewing or checking what was written. On the other hand when a test-taker reads the options of the following question before listening to the questions itself, this is previewing. A test-taker would review or preview as a strategy to double-check their answers to ensure they are correct.

Therefore, as a result of the factor analysis the 49 individual test-taking strategies on the TTSQ were reduced to 15 factors as described above. In the following chapter, these 15 test-taking strategies will be correlated with emotional regulation processes / emotions in order to identify the relationship between them.

Section 2 of the questionnaires focuses on the emotional regulation processes that test-takers used, emotions experienced and behaviors during the test. The following section reports on the emotional regulation processes and emotions identified and on the results of the factor analysis.

5.6.2 EMOTIONAL REGULATION PROCESSES, EMOTIONS AND BEHAVIORS

5.6.2.1 Occurrence of Processes / Emotions / Behaviors

Table (49) shows the means and standard deviations (SD) for each of the TTSQ items related to emotional regulation processes / emotions / behaviors. Test-takers responded using a 6-point rating scale: 6 = I strongly agree; 5 = I agree; 4 = I slightly agree; 3 = I slightly disagree; 2 = I disagree; 1 = I strongly disagree. The items are ordered in the table from the highest to the lowest mean. When interpreting the means of these items, a mean of 4.5 or higher indicated agreement that these processes / emotions / behaviors were experienced. Means ranging from 1.5 to 2.4 indicated disagreement that this feeling was experienced. These processes are shown in bold in the following table:

Table (49): Means & SD of the emotional regulation processes, emotions and behaviors in the TTSQ

#	Emotional Regulation Process / Emotion / Behavior	Mean	SD
AFF28	Thought it was a good idea to have the instructions in both Arabic and English	5.2	1.2
AFF1	Tried hard on this test	5.1	1.17
AFF9	Felt was important to do their best	5.1	1.11
AFF29	Thought it was a good idea to have the writing prompt in both Arabic and English	5.1	1.26
AFF26	Taking the test was a pleasant experience	4.8	1.16
AFF8	Knew what to do during the test	4.7	1.40
AFF13	Felt relieved when the test was over	4.6	1.40
AFF12	Felt a sense of achievement after completing the test	4.3	1.32
AFF24	Felt that the test was interesting	4.3	1.23
AFF2	Did well	4.2	1.11
AFF23	Sure of the correct answer in most questions	4.2	1.24
AFF11	Thought that taking the test was a big challenge	4.1	1.53
AFF10	Had enough time to finish the test	3.8	1.70

Table (49): Means & SD of the emotional regulation processes, emotions and behaviors in the TTSQ (continued)

#	Emotional Regulation Process / Emotion / Behavior	Mean	SD
AFF7	Felt prepared to take this test	3.7	1.56
AFF3	Felt that the test was easy	3.5	1.27
AFF19	Felt tired during the test	3.3	1.72
AFF6	Felt that the test was difficult	3.1	1.49
AFF14	Felt nervous or anxious during the test	3.1	1.62
AFF18	Dread taking tests because they do not show true ability	3.0	1.72
AFF17	Dread taking tests in general	2.9	1.71
AFF27	Felt frustrated because there was not enough time	2.9	1.69
AFF5	Felt that the test was confusing	2.8	1.57
AFF15	Had difficulty in concentrating during the test	2.8	1.57
AFF22	Because of nervousness forgot the things that they usually know	2.8	1.6
AFF25	Felt bored while taking the test	2.5	1.42
AFF16	Kept looking around the testing room during the test	2.1	1.42
AFF20	Gave up because the test was too difficult	2.1	1.43
AFF21	Got tired and so started answering without reading the question	2.1	1.46
AFF4	Was so anxious felt like getting the answer from another person	1.6	1.21

The labels such AFF28 or AFF7, refer to the item number in Section 2 of the TTSQ. From Table (49) it can be seen that the respondents either slightly agreed or slightly disagreed that most of the processes / emotions / behaviors in the questionnaire (18 of the 29 emotional regulation processes, emotions and behaviors) were actually experienced during the test. In other words, the majority of respondents did not admit that these emotions did occur (for example, 41% of the respondents strongly disagreed or disagreed that they felt nervous or anxious during the test while 39% gave a rating of 3 or 4). Thus, when respondents rate these items as 3 or 4 they concede that they experience these processes / emotions to some extent but do not want to take a strong position either way, especially regarding emotional regulation processes associated with negative emotions. It is interesting to note that eleven (11) of these eighteen (18) items reflect negative emotions. In the qualitative data (described later on in this section) the emotions expressed were more or less evenly balanced between positive and negative feelings.

From Table (49) it can be seen the emotions / processes / behaviors that the respondents agreed as having experienced were (means ranged from 4.6 - 5.2):

- thinking it was a good idea to have the instructions in both Arabic and English (AFF28: task-focusing process);
- trying hard on this test (AFF1: behavior);
- thinking it was important to do their best (AFF9: assessment – goal orientation);
- thinking it was a good idea to have the writing prompt in both Arabic and English (AFF29: task-focusing process);
- taking the test was a pleasant experience (AFF26: emotion);
- knowing what to do during the test (AFF8: emotion);
- feeling relieved when the test was over (AFF13: emotion).

The majority of respondents either agreed or strongly agreed that they experienced the processes / emotions / behaviors listed above. All of these are related to a positive affective response and these results are different from those obtained from the think aloud data. From the think aloud data, the most frequently occurring emotional regulation processes / emotions were: feeling that the test or item was difficult, feeling that the test or item was confusing and feeling that the test or item was easy (assessment processes and an emotion and two of the three are related to negative emotions). The reason is probably because emotions are difficult concepts or constructs to define and to make tangible, thus, they are very difficult to quantify. Therefore, differences in data obtained from thinking aloud during a test compared to responding to a questionnaire after completing a test are expected.

From Table (49), respondents disagreed that they experienced the following emotions i.e. they reported that they did not experience these emotions / behaviors:

- kept looking around the testing room during the test (AFF16: behavior);
- gave up because the test was too difficult (AFF20: behavior);
- got tired and so started answering without reading the question (AFF21: emotion);
- was so anxious felt like getting the answer from another person (AFF 4: emotion);

These results are also different from the think aloud data where six of the twelve test-takers reported that they had given up in one or more sections of the test.

5.6.2.2 Qualitative Data on Emotions

Respondents were required to briefly describe their feelings that occurred during the test. 71% of the respondents (351 responses) wrote comments on the open ended question. These responses were coded and the results are shown in Table (50) as follows with the feelings arranged from the most frequently occurring to the least:

Table (50): Frequency counts of coded qualitative data on respondents’ feelings (N=351)

Respondents Feelings	Frequency Count
Positive emotions	122
Negative emotions	105
Test was too long / not enough time	45
Mixed feelings: both positive and negative emotions	40
Test was rewarding and useful especially in finding out their language ability	33
High concentration	23
Neutral feelings	17
Felt challenged in a positive way	13
Positive feelings about test administrators / administration	12
Valued learning English	8
Felt overwhelmed: challenged in a negative way	7
Taking the test was a new experience	7

The positive feelings expressed included delight, joy, feeling happy with the test, feeling relieved, feeling calm, not experiencing anxiety, interesting experience, appreciation, good experience and confidence. Some of the comments made are: “I felt happy with the test because it was very interesting”, “I felt good and there was no anxiety” and “I felt calm, relief and confident.” Eight (8) test-takers commented about the value of learning English. One test-taker stated “I value learning English and every learner must learn English and improve their level” and another commented “I am desperate to learn English...” This reinforces the importance of taking the test since it represents the first step towards learning English. These positive attitudes represent a positive goal orientation and make the test important and relevant to them.

Test-takers also had positive feelings about the test-taking experience itself. Thirteen (13) test-takers felt that the test was a positive challenge making them want to give all they have. Twelve (12) respondents commented very positively about the test administrators. One test-taker commented that "I felt a sense of relief because the test supervisor understood our situation and dealt with us very calmly and was very helpful at the beginning of the test." Several test-takers commented on the manner they were handled and treated during the test. Some comments were "I convey my gratitude and appreciation to all instructors at AUC. I was very happy at this civilized way of dealing with us and the test" and "...this was one of the best tests I have taken. The test administrators dealt with us in a very civilized manner."

The negative emotions experienced were anxiety, fear, worry, lack of concentration, disappointed, feeling upset, feeling confused, frustration because of lack of time, bored, and feeling exhausted. One of the respondents was worried about being placed at a level below his friends. He stated "I prayed that I would do well in this test and that I will be placed in a high level so that I could compete with my friend who have already been placed. Some of them have been placed in level 9 so I hope that I would be at a similar level or higher." Another test-taker commented "I felt tired of fear of the test especially since I had studied English well at university but I did not use it since graduation." As in the think aloud data, several test-takers complained about the sound of the tape not being clear, poor microphones, and the sound quality of the cassette recorder. Several test-takers also complained about the uncomfortable chairs and some stated that this affected their performance on the test. Several test-takers expressed high anxiety, anxiety to some extent and anxiety in some of the sections of the test. One test-taker wrote "I felt confused and bored sometimes and I did not calm down until it was over". Another test-taker stated that "I was anxious during the test this is usual for me in every exam." Seven (7) test-takers felt challenged by the test and commented that it was difficult or that they had a lot to do in a short time.

Several test-takers commented about time pressure and how it led to negative feelings. Some of these comments were: "I answered fast because I feared that time would run out and this made me feel exhausted", "lack of time worried me during the test", "I felt anxious because of the length of the exam, the speed of the listening tape

and not being able to concentrate. Answering fast led to many mistakes” and “time was short and I felt frustrated.”

Test-takers expressed both positive and negative feelings. Some of their comments were: “at the beginning I felt it was a lovely and interesting experience to find out my proficiency level and in the middle I felt physically tired”, “my feelings varied between happiness and feeling upset for reasons I don’t know. Perhaps because I have been away from studying for a long time” and “I felt excited and interesting in answering but sometimes I felt upset when solving some questions.” Several test-takers commented that they liked the test but did not like a particular section. One test-taker stated that it was a wonderful test but she did not like the grammar questions while another did not like the listening section. Several test-takers stated that they started out feeling anxious but it disappeared as time went on: “at the beginning of the test I felt anxious but after the first 5 minutes I felt relief and enjoyment”, “at the beginning I was somewhat anxious and then I was able to calm down” and “I felt relieved because after writing the letter it became clear to me that I can write a letter. I was somewhat confused while listening to the questions but after a while I started listening without anxiety.”

A number of test-takers commented on the usefulness of the test. One test-taker wrote “I felt that this was a useful experience” and another stated “it was an interesting experience because I always wanted to know my level and to increase my abilities by taking tests. I cannot wait to find out the results in order to study what I want.” Several felt very eager to take the test in order to find out their level of language ability. Some test-takers’ comments were “...it provided me with the opportunity to find out my specific level in the language” and “I am not concerned about whether the test is difficult or easy. All I am concerned with is finding out my real proficiency level so that I do not underestimate myself.”

The twenty three (23) respondents who commented about concentration all stated that they concentrated and paid attention during the test. Seventeen (17) responded stating that they had no particular feelings about the test. One test-taker commented that “it was just a placement test to find out my proficiency level and my weakness in order

to study and improve” while another stated “I had no particular feelings and I was concentrating on the test.”

Seven (7) test-takers mentioned that taking this test was a new experience for them. One test-taker wrote that she was anxious in the beginning because she had not been tested in English before. Another test-taker had a similar comment that “this was the first time for me to take an English language test and so I was afraid a little.” It is interesting to note that both of these test-takers are at the intermediate level. Several test-takers commented on the new design of the test as follows: “...I tried a good new type of test I have not taken at school or university before”, “new of its kind” and “It was a new experience. It was enjoyable and different from other tests I have taken. In this test I was more interested and paid more attention.”

The qualitative data yielded similar information to that obtained from the think aloud and TTSQ data regarding positive, negative, neutral and mixed feelings about the test. The new information obtained from the test-takers’ comments was related to test-takers’ feelings about the test administration and administrators and the novelty of taking an English language placement test of this particular design.

The TTSQ also included two open-ended questions designed to further explore test anxiety. One question requested respondents in case they experienced anxiety to indicate in which section of the test they experienced this feeling. There were a total of 314 responses (63%) to this question and the following table shows the frequency count of the various responses.

Table (51): Frequency counts of coded qualitative data on anxiety experienced by test section (N=314)

Test Section where Anxiety was Experienced	Frequency Count
Listening	131
Reading	99
Grammar	22
All sections	5
Writing	4

As shown in Table (51), test-takers experienced most anxiety in the listening section followed by the reading part of the test. This finding is similar to that of the

think aloud data which also revealed the same result. It is not surprising that these test-takers experience anxiety most in the listening section because most Egyptian adults are highly visual learners and thus, they find listening to be very difficult. This is based on the researcher's experience in teaching in this context and on a previous research study conducted on Egyptian adults' learning styles (Boraie et al, 1994).

A follow-up open-ended question was included to further explore different levels of anxiety experienced. The question was: "Did this feeling of anxiety differ from one section of the test to another? How?" Of the 487 test-takers, 298 responded to this question, however, the responses were not very useful. A total of 167 test-takers just commented "yes" while 63 test-takers stated "no" without any further elaboration. There were some interesting individual responses. One test-taker commented "I was normal in all sections except for the last section." Another explained "Yes, I was anxious at the beginning and I gradually became calmer until I reached the end." A test-taker described an opposite experience: "Of course it varied. At the beginning of the test I did not feel it but after reaching the listening I started to experience it and it reached its maximum in the last section in the reading." A test-taker stated "Yes it varied from section to section. For example in the grammar section I did not experience this feeling but I did in the listening and reading sections." Several test-takers indicated that anxiety also varied within one section. For example one test-taker commented "Yes, depending on whether the passages were easy or difficult" and another stated "Yes, it varied when I understood and read the passage completely." A third test-taker commented "Yes, depending on the difficulty and ease of the questions."

Thus, it can be seen that experiencing anxiety is variable and the intensity can increase or decrease significantly depending on the task and the test-taker's ability to perform the particular task. As shown in the LTP model, a blend of emotions is experienced by test-takers in different sections of the test as a result of the continuous assessment processes that occur whenever a test-taker deals with a test task or a particular test question.

5.6.2.3 Results of the Factor Analysis of Emotional Regulation Processes / Emotions / Behaviors

After Varimax rotation, an eight (8) factor solution was obtained which accounted for 60.3% of the total variance in the emotional regulation processes section of the TTSQ as shown in Table (52).

Table (52): Factor analysis for emotional regulation processes / emotions / behaviors

Factor	Label	Eigenvalue	Variance %	Cumulative Variance
Factor 1	Anxiety due to time pressure (emotion)	5.9	11.1	11.1
Factor 2	Goal orientation (assessment)	3.1	7.4	19.3
Factor 3	Reassurance provided by L1 (task-focusing)	1.8	6.3	25.6
Factor 4	Feeling relaxed (emotion)	1.6	5.7	31.3
Factor 5	Behaviors releasing negative emotions (behavior)	1.5	8.7	40.0
Factor 6	Assessment of confidence (assessment)	1.2	6.9	46.9
Factor 7	Fear of taking tests (emotion)	1.1	6.7	53.6
Factor 8	Challenge (emotion)	1.0	6.7	60.3

Tables (53 - 60) describe the questionnaire items that are included in each of the eight (8) factors and the item loadings of each. Six items load on Factor 1 as shown in Table (53):

Table (53): Item loadings on Factor 1

TTSQ Item	Loading
Had enough time to finish the test (AFF10: task focusing)	-0.79
Felt frustrated because there was not enough time (AFF27: emotion)	0.78
Felt that the test was confusing (AFF5: emotion)	0.57
Because of nervousness forgot the things that they usually know (AFF22: emotion)	0.51
Felt that the test was difficult (AFF6: assessment)	0.49
Had difficulty in concentrating during the test (AFF15: emotion)	0.49

Factor 1 has been labeled as “anxiety due to time pressure” because all the items are related to negative feelings and lack of time. AFF27 and AFF10 which negatively

correlates with the rest of the items indicate that this factor involves time pressure where the test-taker feels that he / she does not have enough time to complete the test. Confusion, feeling that the test was difficult and forgetting things because of nervousness all relate to anxiety. It is interesting to note that “difficulty in concentrating” loads on this factor. This result is similar to a previous study on motivation conducted by the researcher in the same context as this study (Schmidt et al, 1996) where in a factor analysis difficulty in concentration also loaded on anxiety. Anxiety is related to the intrusion of unwelcome thoughts leading to difficulty in concentrating. Although the items loading on this factor include a task-focusing and an assessment process, however, the interpretation of Factor 1 suggests it is an emotion related to lack of time.

Factor 2 consists of four questionnaire items as follows:

Table (54): Item loadings on Factor 2

TTSQ Item	Loading
Felt was important to do their best (AFF9: assessment)	0.80
Tried hard on this test (AFF1: behavior)	0.69
Knew what to do during the test (AFF8: task-focusing)	0.56
Felt prepared to take this test (AFF7: emotion)	0.53

Factor 2 was labeled “goal orientation”. This factor includes a positive goal orientation of feeling that it was important to exert one’s maximum effort on the test. Here the test-taker felt as if it was a job well done. A high score on this factor would indicate that a test-taker had a positive goal orientation and exerted all his / her efforts to do well on this test. He / she did what had to be done and there was no feeling of guilt or any other negative emotion. This process can be classified as an assessment process.

The two items that load on Factor 3 are shown in Table (55):

Table (55): Item loadings on Factor 3

TTSQ Item	Loading
-----------	---------

Thought it was a good idea to have the instructions in both Arabic and English (AFF28: task focusing)	0.92
Thought it was a good idea to have the writing prompt in both Arabic and English (AFF29: task focusing)	0.91

Factor 3 was labeled as “reassurance provided by L1.” The two items that load on this factor are related to the presence of Arabic in the instructions or in the description of the test task (writing prompt). They are both associated with a positive emotion and from the think aloud data, the test-takers expressed approval or felt pleased at finding directions and the prompt in Arabic. This process is classified as task-focusing.

Factor 4 consists of two items as shown in Table (56):

Table (56): Item loadings on Factor 4

TTSQ Item	Loading
Felt that the test was interesting (AFF24: emotion)	0.80
Taking the test was a pleasant experience (AFF26: emotion)	0.76

Factor 4 is very clearly associated with positive feelings where the test-taker is feeling relaxed and finds the test to be both pleasant and interesting no tens. This factor is labeled as “feeling relaxed” and is classified as an emotion.

Five items load on Factor 5 as shown in Table (57):

Table (57): Item loadings on Factor 5

TTSQ Item	Loading
Kept looking around the testing room during the test (AFF16: behavior)	0.67
Felt bored while taking the test (AFF25: emotion)	0.64
Got tired and so started answering without reading the question (AFF21: emotion)	0.54
Gave up because the test was too difficult (AFF20: behavior)	0.53
Was so anxious felt like getting the answer from another person (AFF4: emotion)	0.50

The three items related to behavioral effects that occur as a result of negative emotions load on this factor. These behaviors are giving up, looking around the testing room and the feeling of wanting to cheat. It is interesting to find that boredom and feeling tired are also associated with these behaviors. Thus, when experiencing negative emotions such as anxiety a test-taker may just lose interest and express boredom, wanting to decrease the emotional stress. Feeling tired is also a similar manifestation of the efforts exerted by the test-taker to deal with negative emotions. Factor 5 is labeled as “behaviors releasing negative emotions” and it is a behavioral process associated with negative feelings.

Factor 6 consists of three questionnaire items as follows:

Table (58): Item loadings on Factor 6

TTSQ Item	Loading
Sure of the correct answer in most questions (AFF23: task-focusing)	0.74
Did well (AFF2: assessment)	0.61
Thought that the test was easy (AFF3: assessment)	0.52

Factor 6 is labeled as “assessment of confidence.” Here the test-taker is confident in his / her ability and felt he / she did well (positive self assessment). The test-taker was sure that he / she answered correctly feeling that the test was easy. A low score on this factor would indicate a test-taker of low confidence who would negatively assess his / her ability. The factor is classified as an assessment process.

Two items load on Factor 7 as shown in Table (59):

Table (59): Item loadings on Factor 7

TTSQ Item	Loading
Dread taking tests because they do not show true ability (AFF18: emotion)	0.78
Dread taking tests in general (AFF17: emotion)	0.74

Factor 7 is labeled as “fear of taking tests” and both items are clearly related to this highly negative emotion towards taking tests. This factor is related to previous test-taking experiences. As shown in the think aloud data, previous test-taking experiences

do have an impact on testing situations later on. When faced with a test, the test-taker may recall positive or negative test-taking experiences. If the test-taking experiences were highly negative, then recalling them would lead to negative emotions even before the test-taker receives the test paper. As one test-taker noted in the qualitative data in the TTSQ “it brought back the negative experiences.” Factor 7 is classified as an emotion.

Factor 8 consists of three questionnaire items as follows:

Table (60): Item loadings on Factor 8

TTSQ Item	Loading
Felt relieved when the test was over (AFF13: emotion)	0.83
Felt a sense of achievement after completing the test (AFF12: emotion)	0.82
Felt that taking the test was a big challenge (AFF11: emotion)	0.45

Factor 8 was labeled as “challenge.” This factor is associated with the positive feeling of being challenged during the test accompanied by relief and a sense of achievement at the end. According to Csikszentmihalyi & Nakamura (1989) affect will be highest when challenge and skills (the ability of the test-taker) are perceived to be about equal and when both are high. When the challenge of a task is high and skills are low then anxiety emerges and when challenge and skill are both low the outcome is lack of interest or even boredom. Thus, a high score on Factor 8 would indicate that the test-taker perceives the task as a challenging task within his / her ability. A low score would indicate the test was beyond his / her skills and the test-taker will not experience relief of a sense of achievement at the end of the test. This factor is classified as an emotion.

Therefore, the factor analysis identified two positive emotions, two negative emotions, one task-focusing process, two assessment processes and a behavioral process. This validates the various components of emotional regulation and test-taking behavior as postulated in the LTP model.

5.7 TEST-TAKING STRATEGIES AND EMOTIONAL REGULATION PROCESSES IDENTIFIED

Most of the test-taking strategies identified in this study from the think aloud matched those found in the literature, specifically those mentioned by Nevo (1989), Anderson et al (1991) and Cohen (1998). The seventeen strategies that were identified from the think aloud data and not from the literature were mostly related to the use of L1 and were related to the specific design of the test which included directions and the writing prompt written in Arabic. In the literature reviewed, translation was mentioned as a strategy used to comprehend a word or a phrase. However, in this study from the think aloud data most test-takers translated the questions and some or all of the options throughout the test and not just words or phrases. It is interesting to note that in the questionnaire data, translation did not emerge as a high frequency strategy and this may be because test-takers may be translating automatically without being really conscious of the application of this strategy.

Another strategy used by test-takers while taking a reading test and not reported in the test-taking strategies literature was the sounding out of a word in an attempt to find its meaning. The test-taker was probably trying to trigger the phonological form of the word from memory in an attempt to recall its meaning. This strategy may be related to the test-takers L1 background which has a completely different script from English and where the sound of a word and its meaning are linked in a complex system using diacritics.

As mentioned by Cohen (1998) in several instances, test-takers' correct responses did not necessarily reflect their knowledge or language ability such as using the strategy of guessing and getting the question right. Nevo (1989) had suggested that test-taking strategies could be classified as either contributing to the correct response or not contributing to the correct answer. The findings showed that strategies could not be classified as either contributory or non-contributory since in many instances in the think aloud data, it was found that the same test-taking strategy used by one test-taker could be both contributory and non-contributory.

In this study, similar to the findings by Wijgh (1996) and Weir et al (2000), test-takers do not necessarily use the strategies intended by the test designers. For example,

in the listening task in Parts B and C, the test designers intended to test learners' ability to comprehend a spoken sentence. Some of the test-takers would use the strategy of focusing on one or two words they hear and then look at the options attempting to locate the word they just heard. Thus, they were not really listening at the sentence level, they only focused on individual words.

There have been no studies on emotional regulation while taking language tests and thus, no comparisons of the findings of this study with the literature on language testing can be made. However, the findings of this study do support to a large extent the findings of studies conducted in education in general. Schutz & Davis (2000) and Schutz et al (2002) had proposed a model of emotional regulation during test-taking that included cognitive-appraising, task-focusing, and emotion-focusing processes and emotions with test-taking strategies as a separate component. The findings of this study supported the literature showing that test-takers use cognitive-appraising and task-focusing processes with emotions emerging and test-taking strategies as a separate component. However, the findings could not identify any emotion-focusing processes used while taking a test.

On comparing the emotions that emerged from the data with the emotions identified from the literature, overlap was found. Almost all the emotions that were found in the data had been previously identified but with different labels in some cases. For example, self-satisfaction was not mentioned in the literature review conducted by Pekrun et al (2002) on emotions in learning and achievement, however it is related to other emotions listed such as joy, enthusiasm and pride. Because the literature was not focused only on emotions in test-taking contexts there were many positive emotions listed. There were far fewer positive emotions compared to negative emotions emerging from the data.

The test-taking strategies and the emotional regulation processes identified are summarized in the following section.

5.8 SUMMARY OF FINDINGS AND OUTCOMES ON RESEARCH QUESTION 2

The focus of Research Question 2 is on identifying the test-taking strategies and emotional regulation processes used and emotions experienced by Egyptian adult EFL learners while taking an English language placement test. The findings are based on both the think aloud and TTSQ data.

- Findings regarding high frequency test-taking strategies differed to some extent between the think aloud and TTSQ data. The TTSQ data indicated that high frequency test-taking strategies are related to test-takers' attempts at clearly understanding what is required from the task. In the think aloud data two of the four high frequency strategies (which appeared as low frequency strategies in the TTSQ data) are "guessing" and "translation". The other two strategies are similar to those obtained from the TTSQ data: "reading the question and options" and "reading instructions carefully".
- Findings regarding low frequency test-taking strategies were similar from both the think aloud and TTSQ data. The six low frequency test-taking strategies were: "postponing dealing with a question", "selecting an option that was longer / shorter than others", "reviewing / surveying the test", "watching to see when other students finish the test", "reading the prompt in Arabic only" and "reading the question first before reading the passage."
- Findings regarding high frequency emotional regulation processes and emotions differed between the think aloud and TTSQ data. The high frequency emotional regulation processes / emotions / behaviors from the TTSQ data are all related to positive emotions. These are: thinking it is a "good idea to have the instructions and the writing prompt in both Arabic and English", "trying hard on the test", "feeling it was important to do their best", "knowing what to do", "feeling that taking the test was a pleasant experience" and "feeling relieved when the test was over." In the think aloud data the three high frequency emotional regulation processes were two assessment processes and an emotion, however, two of the three are related to negative emotions: feeling the test / item was easy, difficult and confusing.

- Findings regarding the low frequency emotional regulation processes and emotions also differed between the think aloud and TTSQ data. From the TTSQ data, the low frequency emotional regulation processes / emotions / behaviors are: “looking around the room during the test”, “giving up”, “getting tired and started to answer without reading” and was “so anxious felt like getting the answer from another person”. From the think aloud data the low frequency emotional regulation processes were all related to enjoying the test and having a goal orientation. In the think aloud data six of the twelve test-takers reported giving up in one or more sections of the test.
- From the qualitative data obtained from the TTSQ, five new test-taking strategies were identified: “selecting an answer based on a word heard in the utterance”, “stopping to take a break”, “recalling information about memorized words”, “choosing the middle option when answer was not known” and “refusing to answer a question when not convinced with the options or not sure of the answer.” Two new emotional regulation processes / emotions were identified: “positive feelings related to the test administration process and the test administrators themselves” and the “novelty of taking an English language placement test of this particular design”.
- From both the think aloud and TTSQ qualitative data it was found that anxiety was experienced most in the listening section followed by the reading. It was also found that the intensity of anxiety varied across sections and within one section in different items.
- Using factor analysis, the TTSQ data was reduced to 15 test-taking strategies and 8 emotional regulation processes, emotions and behaviors as shown in Table (61):

Table (61): Test-taking strategies and emotion regulation processes / emotions / behaviors experienced by Egyptian adult EFL learners

Test-taking Strategies	Emotional Regulation Processes / Emotions / Behaviors
Making use of clues	Anxiety due to time pressure
Dealing with unknown texts	Goal orientation
Using L1 to deal with L2	Reassurance provided by L1
Leaving blank responses	Feeling relaxed
Managing time	Behaviors releasing negative emotions
Using knowledge of grammar rules	Assessment of confidence
Dealing with unknown words	Fear of taking tests
Rereading text	Challenge
Handling MCQ options	
Recognizing correct answers without hesitation	
Elimination	
Rereading questions	
Ensuring understanding	
Reading the prompt in two languages: L1/L2/L1	
Previewing / reviewing	

- Based on the think aloud data, it was found that the LTP model is supported with some components revised to accurately describe the processes involved in a test-taking situation. The factor analysis of the quantitative data also supported the various components of the model. Test-taking strategies are no longer classified as contributory and non-contributory. Emotional regulation consists of assessment and task focusing processes and emotions. In case of a negative affective response, negative emotions emerge and task-focusing processes are used along with test-taking strategies to complete the task; and for positive and neutral emotions the test-taker just carries on what he / she is doing using task-focusing processes and test-taking strategies to complete the task. The test-taker's behavior during the test-taking process is the outcome of an initial assessment followed by the interplay between the selection and adaptation of test-taking strategies, the emotions that emerge and of task-focusing processes.

CHAPTER 6:

RESULTS & FINDINGS FOR RESEARCH QUESTION 3

6.1 RESEARCH QUESTION 3

What is the relationship between test-taking strategies, emotional regulation and test performance for Egyptian adult EFL learners in a specific context when taking an English language placement test? Further questions that focus on the relationship between test-taking strategies, emotional regulation and test performance are:

- 3.1 Do test-taking strategies vary across different levels of test performance?
- 3.2 If so, how do test-taking strategies differ across different levels of test performance?
- 3.3 Does emotional regulation vary across different level of test performance?
- 3.4 If so, how does emotional regulation differ across different levels of test performance?
- 3.5 Is there a relationship between test-taking strategies selected and emotional regulation?
- 3.6 If so, does the relationship between test-taking strategies selected and emotional regulation differ across different levels of test performance?

Research Question 3 will be addressed based on the findings from the think aloud data analyzed in detail in Chapter 5 and further analysis of the data obtained from administering the Test-taking Strategies Questionnaire (TTSQ). The TTSQ is described in detail in the Research Design Chapter.

6.2 PROFICIENCY PROFILE OF RESPONDENTS

Table (62) shows the distribution of the 497 respondents by proficiency level.

Table (62): Distribution of respondents by proficiency level (N=497)

Proficiency Level	Frequency Count	Percentage
Novice	72	14.5%
Elementary	219	44.1%
Intermediate	164	33%
Advanced	42	8.5%

Table (62) shows that the majority of test-takers were placed in the elementary (44%) and intermediate (33%) levels. This distribution is typical of the CACE population studying English. The majority of classes are at the elementary and intermediate levels with fewer classes at the novice level and even fewer at the advanced level.

The TTSQ data are analyzed in terms of this distribution by proficiency level in order to relate test-taking strategies and emotional regulation to test performance. In this study, proficiency level i.e. the test-taker's placement on the basis of the test score, is equal to the test-taker's test performance. It should be noted here that a test-taker's actual proficiency level or language ability may not be equal to that of the test result. If a test-taker experiences high anxiety, he / she may perform very poorly on the test and thus, the resulting proficiency level may be far lower than the test-taker's actual ability. Measuring this actual ability is very difficult or may not be even possible. If an individual experiences emotional distress in any test situation then it is not possible to block this emotion or stop its occurrence. Thus, we cannot measure true ability separately from emotion for this individual. The focus of the study is to explore the relationship between test-taking strategies, emotional regulation and test performance to find out whether test-takers at different proficiency levels use different test-taking strategies and emotional regulation, whether the proficiency level of test-takers can be predicted by utilization of particular combinations of strategies and emotional regulation processes, whether the use of different strategies is related to emotional regulation and whether the relationship between test-taking strategies and emotional regulation, if any, varies by proficiency level.

6.3 DO TEST-TAKING STRATEGIES VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?

In order to address this question, the means and SD of the respondents' scores on the 15 test-taking strategies identified in Research Question 2 were obtained at each of the four proficiency levels as shown in Table (63). The standard deviations are indicated in brackets in the table.

Table (63): Means & SD of test-taking strategies by proficiency level

Test-taking Strategy	Proficiency Level			
	Novice	Elem.	Int.	Advanced
Making use of clues	2.3 (0.62)	2.3 (0.50)	2.2 (0.46)	2.0 (0.51)
Dealing with unknown texts	1.8 (0.35)	1.8 (0.25)	1.7 (0.18)	1.7 (0.15)
Using L1 to deal with L2	2.0 (0.44)	2.0 (0.38)	1.8 (0.41)	1.5 (0.24)
Leaving blank responses	2.1 (0.65)	1.8 (0.64)	1.5 (0.50)	1.1 (0.35)
Managing time	1.6 (0.79)	1.7 (0.76)	1.8 (0.82)	1.5 (0.71)
Using knowledge of grammar rules	2.1 (0.69)	2.4 (0.61)	2.4 (0.59)	2.3 (0.67)
Dealing with unknown words	2.2 (0.66)	2.2 (0.58)	2.2 (0.53)	1.9 (0.55)
Rereading text	2.0 (0.64)	2.0 (0.52)	2.0 (0.51)	2.0 (0.42)
Handling MCQ options	1.8 (0.65)	1.6 (0.51)	1.6 (0.54)	1.7 (0.52)
Recognizing correct answers without hesitation	1.9 (0.80)	1.9 (0.77)	1.8 (0.82)	1.7 (0.77)
Elimination	1.9 (0.58)	2.0 (0.55)	2.0 (0.44)	2.0 (0.47)
Rereading questions	2.3 (0.55)	2.1 (0.53)	1.9 (0.47)	1.8 (0.47)
Ensuring understanding	2.4 (0.54)	2.5 (0.56)	2.6 (0.46)	2.6 (0.42)
Reading the prompt in two languages: L1/L2/L1	2.0 (0.53)	1.9 (0.50)	1.8 (0.48)	1.7 (0.51)
Previewing / reviewing	1.8 (0.59)	1.8 (0.51)	1.9 (0.54)	2.0 (0.44)

In order find out whether these test-taking strategies differed across proficiency levels, fifteen one-way ANOVAs were computed one for each strategy. The results of the ANOVAs were found to be significant for five (5) test-taking strategies ($p<0.002$):

- Using L1 to deal with L2 $F(3, 493) = 24.2$
- Leaving blank responses $F(3, 493) = 36.4$
- Using knowledge of grammar rules $F(3, 493) = 5.9$
- Rereading questions $F(3, 493) = 12.1$
- Reading the prompt in two languages $F(3, 493) = 9.1$

Therefore, the answer to research question 3.1 is that test-taking strategies vary to some extent across different levels of test performance. Five of the fifteen test strategies were used differently at different levels of proficiency as shown in the following section.

6.4 IF SO, HOW DO TEST-TAKING STRATEGIES VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?

The post hoc comparisons using the Scheffe test and examining the means revealed that the advanced group used the strategy of using L1 to deal with L2 significantly lower than the other three groups. This is understandable because the advanced level test-takers would depend less on L1 (Arabic) than the other groups who need the help and support of Arabic.

Further post hoc comparisons and examination of the means revealed that the advanced group used the strategy of leaving blank answers significantly lower than the other three groups. This is expected since advanced level test-taker would most likely not leave any answers blank and would probably be able to complete the test within the assigned time. In fact, test-takers in the novice group may intentionally leave items blank and this was illustrated in the qualitative data where a novice stated that “I was afraid that I would not be able to show my true level therefore, I did not mark many of the answers.”

It was found that the intermediate group used knowledge of grammar rules the most compared to the other three groups. Perhaps at this proficiency level, learners are not yet quite fluent compared to the advanced group and definitely have more grammar knowledge compared to the novice and elementary groups. Therefore, they may depend on grammar more than other proficiency groups when taking a test.

As expected the novice group tended to use the strategy of rereading questions more than the other three groups. Thus, the advanced group used this strategy the least. A similar pattern was found for the strategy of reading the prompt in two languages: L1/L2/L1. It was found that the novice group used this strategy the most which means they would “read the prompt first in Arabic and then in English” more often compared to the other groups. Thus, advanced group test-takers would more likely read the prompt in English first and then in Arabic.

Thus, it was found that the novice and advanced groups differed most in strategy use in four of the five test-taking strategies that were identified as being significantly different across the proficiency levels. In only one test-taking strategy the intermediate group used the strategy of “using knowledge of grammar rules” differently.

6.5 DOES EMOTIONAL REGULATION VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?

Table (64) shows the means and SD of the respondents’ scores on the 8 emotional regulation processes / emotions / behavior identified in Research Question 2 at each of the four proficiency levels. The standard deviations are indicated in brackets in the table.

Table (64): Means & SD of emotional regulation processes by proficiency level

Emotional Regulation Process	Proficiency Level			
	Novice	Elem.	Int.	Advanced
Anxiety due to time pressure	3.3 (0.80)	3.2 (0.79)	2.9 (0.77)	2.4 (0.66)
Goal orientation	4.2 (1.13)	4.6 (0.87)	4.9 (0.75)	4.9 (0.87)
Reassurance provided by L1	5.6 (0.88)	5.4 (0.99)	4.9 (1.24)	4.5 (1.52)
Feeling relaxed	4.3 (1.2)	4.6 (0.98)	4.7 (1.0)	4.5 (1/04)
Behaviors releasing negative emotions	2.6 (0.92)	2.2 (0.96)	1.7 (0.72)	1.6 (0.46)
Assessment of confidence	3.1 (0.92)	3.8 (0.80)	4.4 (0.68)	4.8 (0.75)
Fear of taking tests	3.5 (1.67)	3.0 (1.46)	2.8 (1.43)	2.5 (1.03)
Challenge	4.1 (1.26)	4.3 (1.10)	4.6 (0.99)	4.1 (1.23)

In order find out whether these emotional regulation processes differed across proficiency levels, eight one-way ANOVAs were computed one for each process. The results of the ANOVAs were found to be significant for five (5) emotional regulation processes ($p<0.002$):

• Anxiety due to time pressure	$F(3, 493) = 20.5$
• Goal orientation	$F(3, 493) = 12.4$
• Reassurance provided by L1	$F(3, 493) = 12.9$
• Behaviors releasing negative emotions	$F(3, 493) = 26.5$
• Assessment of confidence	$F(3, 493) = 58.4$

Therefore, the answer to research question 3.2 is that emotional regulation processes do vary to a great extent across different levels of test performance. Five of the eight emotional regulation processes were used differently at different levels of proficiency as shown in the following section.

6.6 IF SO, HOW DOES EMOTIONAL REGULATON VARY ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?

The post hoc comparisons using the Scheffe test revealed that the advanced group experienced significantly lower “anxiety due to time pressure” compared to the other groups. It was also found that “goal orientation” was lowest for the novice group. As expected “reassurance provided by L1” was highest for the novice group since they probably need the most help and support. The advanced group reported “behaviors releasing negative emotions” significantly less than the other groups. “Assessment of confidence” was also the highest for the advanced group which is not surprising since they have the highest ability compared to the other groups.

Thus, it can be seen that emotional regulation does vary for the advanced and novice proficiency levels. The advanced test-takers experienced the highest confidence, the least anxiety and the least frequent use of behaviors used to cope with negative emotions. Novice test-takers experienced the least goal orientation and the most reassurance provided by the presence of Arabic instructions and the Arabic writing prompt in the test.

Therefore, test-taking strategies vary to some extent across the different proficiency levels while emotional regulation does vary to a great extent. The following section explores the interplay between emotional regulation and test-taking strategies.

6.7 IS THERE A RELATIONSHIP BETWEEN TEST-TAKING STRATEGIES SELECTED AND EMOTIONAL REGULATION?

The fifteen test-taking strategies and the eight emotional regulation processes / emotions / behaviors were correlated to identify the possible relationships. Pearson product-moment correlation coefficients were computed and Table (65) shows the significant correlations.

Table (65): Significant correlations between test-taking strategies and emotional regulation processes / emotions / behaviors

Correlations between ...	Correlation Coefficient	P <
Making use of clues and anxiety due to time pressure	0.22	0.05
Dealing with unknown texts and anxiety due to time pressure	0.27	0.01
Using L1 to deal with L2 and anxiety due to time pressure	0.32	0.01
Leaving blank responses and anxiety due to time pressure	0.26	0.01
Elimination and anxiety due to time pressure	0.21	0.05
Rereading questions and anxiety due to time pressure	0.24	0.02
Ensuring understanding and goal orientation	0.22	0.05
Using L1 to deal with L2 and reassurance provided by L1	0.23	0.05
Rereading questions and reassurance provided by L1	0.22	0.05
Reading prompt in 2 languages and reassurance provided by L1	0.21	0.05
Dealing with unknown texts and behaviors releasing negative emotions	0.28	0.01
Leaving blank responses and behaviors releasing negative emotions	0.27	0.01
Using knowledge of grammar rules and assessment of confidence	0.21	0.05
Previewing / reviewing and assessment of confidence	0.24	0.02
Dealing with unknown texts and fear of taking tests	0.20	0.05

Table (65) shows fifteen (15) significant correlations indicating that there is some relationship between test-taking strategies selected and emotional regulation processes / emotions / behaviors. It is interesting to note that there are no significant correlations between test-taking strategies and positive emotions. Test-taking strategies

correlate with task-focusing processes, assessment processes, negative emotions, and behaviors. This is related to the findings from the think aloud data where it was difficult to identify emotional regulation processes associated with positive emotions. There are no significant correlations between emotional regulation and five test-taking strategies: “managing time”, “dealing with unknown words”, “rereading text”, “handling MCQ options” and “recognizing correct answers without hesitation.”

Six test-taking strategies correlate with anxiety. When test-takers experience “anxiety due to time pressure”, they are likely to attempt to search for clues to find the right answer, translate (use L1 to deal with L2), leave blank responses, eliminate options and reread questions. It is interesting to note that “dealing with unknown texts” correlates with anxiety. This is not surprising because as previously mentioned (see Chapter) the task of reading an unseen text is a difficult task for many test-takers and thus, it is expected to correlate with anxiety. Dealing with unknown texts is also related to behaviors releasing negative emotions which include giving up, feeling tired, bored or looking around the testing room. It seems that for some test-takers dealing with unknown texts was a very negative experience either during this test or in previous test because it correlated with fear of taking tests.

Goal orientation correlates with the test-taking strategy of ensuring answering when responding carefully. This strategy is characteristic of a test-taker who knows what he / she is doing. The emotion of reassurance provided by L1 naturally correlates with the test-taking strategies of using L1 to deal with L2 (translation, formulating ideas in Arabic,...) and reading the writing prompt first in Arabic and then in English or vice versa. Interestingly rereading questions correlated with reassurance provided by L1. Perhaps rereading here is associated with the test-taker mentally translating the question in an attempt to understand and respond to the task.

The test-taking strategy of “leaving blank responses” is correlated with behaviors releasing negative emotions. Thus, test-takers experiencing negative emotions and this behavioral process tend to leave questions blank. Assessment of confidence correlates with the test-taking strategies of using knowledge of grammar rules. In an education context where the main teaching methodology used was the grammar translation approach, it is expected that adult learners who are a product of this system would gain

confidence by using knowledge of grammar during a test. Previewing / reviewing is a cognitively demanding test-taking strategy where the test-taker reviews what he/ she has written or in the listening reads ahead the options of the following item before listening to the question. Thus, a test-taker must have the confidence to use this strategy.

Based on the results of the correlations, it was established that there is some relationship between test-taking strategies and emotional regulation processes / emotions / behaviors. In order to investigate this relationship further, it was important to identify specific sub-groups among the respondents and then examine similarities and differences. It was difficult to classify respondents on the basis of strategy use, however, it was possible to identify specific sub-groups of respondents according to emotional regulation. The researcher decided to select two emotional regulation processes: one that involved positive emotions and the other that resulted in a negative affective state to be used to identify sub-groups of respondents. Anxiety due to time pressure was selected since it emerged as Factor 1 in the emotional regulation processes. A score of 4.5 or greater on anxiety indicates that the test-taker reported experiencing high anxiety or high confidence during the test while a score of less than 2.5 indicates low anxiety (these means are set in relation to the 6-point scale in section 2 of the TTSQ). Using these means test-takers were identified at the high and low ends of “goal orientation”, “feeling relaxed”, “assessment of confidence” and “challenge”. In order to carry out a discriminant analysis, each group (high and low) should include at least 20 test-takers. Furthermore, care should be taken when dealing with unequal groups. On examining the number of test-takers at each end of these three variables, “assessment of confidence” (Factor 6) was selected for analysis (there were only 12 test-takers with a low goal orientation and 16 test-takers who were low on feeling relaxed). The scores of each test-taker on Factor 1 and Factor 6 were examined with the result that the test-takers were divided into a high anxiety or a low anxiety group and a high confidence or low confidence group. In general, there were different test-takers in these four groups, however, there were some test-takers who appeared in one of the high / low anxiety groups and also in one of the high / low confidence groups. Test-takers whose scores fell in the range of 1.5 – 4.4 were not classified in any of the groups. These test-takers reported experiencing anxiety or confidence to some extent.

Of the 497 respondents, there were 31 high anxiety (HA) test-takers and 131 low anxiety (LA) test-takers. Table (66) shows the distribution of the test-takers in each of the HA and LA groups by proficiency level.

Table (66): Distribution of high anxiety and low anxiety test-takers by proficiency level

Proficiency Level	High Anxiety Group	Low Anxiety Group
Novice	4 (13%)	8 (6%)
Elementary	18 (58%)	38 (29%)
Intermediate	9 (29%)	58 (44%)
Advanced	--	27 (21%)
Total	31	131

As seen from the data in Table (66) and similar to previous results, there no advanced test-takers in the high anxiety group with the majority in the elementary group. The majority (44%) of the test-takers in the low anxiety group were at the intermediate level.

Of the 497 respondents, there were 31 high confidence (HC) test-takers and 131 low confidence (LC) test-takers. Table (67) shows the distribution of the test-takers in each of the HC and LC groups by proficiency level.

Table (67): Distribution of high confidence and low confidence test-takers by proficiency level

Proficiency Level	High Confidence Group	Low Confidence Group
Novice	5 (4%)	18 (58%)
Elementary	36 (26%)	13(42%)
Intermediate	70 (52%)	--
Advanced	24 (18%)	--
Total	135	31

In Table (67) it can be seen that there are no intermediate or advanced test-takers in the low confidence group while the majority (52%) of the high confidence group are test-takers at the intermediate level.

6.7.1 Discriminant Analysis by Anxiety

In order to determine which test-taking strategies discriminate between high and low anxiety test-takers and high and low confidence test-takers, a step-wise Discriminant Analysis (DA) was computed for each of the emotion and assessment process (anxiety and confidence) and the discriminating variables were the 15 test-taking strategies. For more details on the statistical technique of DA, refer to Klecka (1980). The DA results are shown in Table (68):

Table (68): Results of discriminant analysis by anxiety

Discriminant Function	Eigenvalue	Canonical R	Wilks' lambda	F-test	Df	P
Anxiety	0.53	0.59	0.65	8.93	9, 152	0.00

The results of the Wilks' Lambda statistic indicated that the overall test had a significant results at $p = 0.00$ level. The discriminatory power of the discriminatory function (the size of the discriminatory function is used to identify the importance of the independent variable) is expressed by the eigenvalue of 0.53 and the degree of association between the groups and the discriminant scores is expressed by the canonical correlation of 0.59. Thus, there is a significant difference between the group means, however there is a large amount of variance (65% of the variance) not accounted for by the test-taking strategies. This is not surprising since factors other than test-taking strategies such as language ability would account for difference in emotional regulation processes and emotions.

Since the objective is to minimize the number of misclassifications, the Jackknifed classification matrix (a matrix used to assess the performance of DA), shown in Table (69) is also examined:

Table (69): Jackknifed classification matrix for high / low anxiety groups

	High	Low	% Correct
High	24	7	77%
Low	26	105	80%
Total	50	112	80%

Because of unequal group size, it is important to examine the classifications closely. Large groups have a disproportionally higher chance of classification. However, this does not seem to be the case as shown in Table (69). Table (69) shows that of the 31 cases in the high anxiety group, 24 (77%) were predicted correctly to be members of the high anxiety group while 7 were misclassified to the low anxiety group. Similarly, 105 out of the 131 cases (80%) in the low anxiety group were identified correctly and 26 were assigned incorrectly to the high anxiety group. The overall percentage of cases classified correctly is good: 80%.

Table (70) shows the canonical discriminant coefficients for each of the predictors that contributed to the classification between the two groups.

Table (70): Standardized canonical discriminant coefficients for anxiety

Discriminating Variable	Loading
Making use of clues	0.22
Dealing with unknown texts	0.24
Using L1 to deal with L2	0.35
Leaving blank responses	0.34
Handling MCQ options	0.21
Recognizing correct answers without hesitation	0.34
Rereading questions	0.28
Ensuring understanding	-0.34
Reading the prompt in two languages: L1/L2/L1	0.27

In general any variable with a loading of ± 0.3 or more is considered to be important in defining the discriminant function (anxiety). Thus, the four test-taking strategies “using L1 to deal with L2”, “leaving blank responses”, “recognizing correct answers without hesitation” and “ensuring understanding” are those that discriminated between the high / low anxiety groups. On examining the positive and negative values of the coefficients, “ensuring understanding” is the strategy most likely used by the low anxiety group whereas the other three strategies are used by the high anxiety group. The use of the strategy of “recognizing correct answers without hesitation” by the high anxiety group is interesting. This means that it is an impulsive test-taker who just seizes on to the first answer that seems correct without checking further. The DA analysis showed that test-takers with high anxiety do use significantly different test-taking

strategies from the low anxiety group. These results agree with reported in the literature on test anxiety. Schutz et al (2002:337) state that in previous research it was found that “there is a tendency for high test-anxious students to report using more strategies (task- and emotion-focused) than low test-anxious students...”

6.7.2 Discriminant Analysis by Confidence

The DA results were similar to that of anxiety and are shown in Table (71):

Table (71): Results of discriminant analysis by confidence

Discriminant Function	Eigenvalue	Canonical R	Wilks' lamba	F-test	Df	P
Confidence	0.56	0.60	0.64	12.70	7, 158	0.00

The results of the Wilks' Lambda statistic indicated that the overall test had a significant results at $p = 0.00$ level. The discriminatory power of the discriminatory function is expressed by the eigenvalue of 0.56 and the degree of association between the groups and the discriminant scores is expressed by the canonical correlation of 0.60. Thus, there is a significant difference between the group means, however there is a large amount of variance (64% of the variance) not accounted for by the test-taking strategies. Once again, this is not surprising since factors other than test-taking strategies such as language ability would account for difference in emotional regulation processes.

The Jackknifed classifiation matrix is shown in Table (72) as follows:

Table (72): Jackknifed classification matrix for high / low confidence groups

	High	Low	% Correct
High	112	23	83%
Low	6	25	81%
Total	118	48	83%

Because of unequal group size, it is important to examine the classifications closely. Table (72) shows that of the 135 cases in the high confidence group, 112 (83%) were predicted correctly to be members of the high confidence group while 23 were misclassified to the low confidence group. Similarly, 25 out of the 31 cases (81%) in the low confidence group were identified correctly and 6 were assigned incorrectly to the high confidence group. The overall percentage of cases classified correctly is good: 83%.

Table (73) shows the canonical discriminant coefficients for each of the predictors that contributed to the classification between the two groups.

Table (73): Standardized canonical discriminant coefficients for confidence

Discriminating Variable	Loading
Making use of clues	0.38
Leaving blank responses	0.47
Using knowledge of grammar rules	-0.50
Elimination	0.26
Rereading questions	0.38
Ensuring understanding	-0.49
Previewing / reviewing	-0.51

The one variable with a loading of less than 0.3 (“elimination”) is not included in the interpretation. Thus, the six test-taking strategies “making use of clues”, “leaving blank responses”, “using knowledge of grammar rules”, “rereading questions”, “ensuring understanding” and “previewing / reviewing” are those that discriminated between the high / low confidence groups. On examining the positive and negative values of the coefficients, it can be seen that “using knowledge of grammar rules”, “ensuring understanding” and “previewing / reviewing” are most likely used by the high confidence group while the other three strategies are used by the low confidence group. The DA analysis showed that test-takers with a high assessment of confidence do use significantly different test-taking strategies from those with low confidence.

Therefore, both the correlations and the discriminant analysis showed a definite relationship between emotional regulation processes, emotions and test-taking strategies.

Test-takers experiencing different emotions and regulation processes do use different test-taking strategies throughout the test.

6.8 IF SO, DOES THE RELATIONSHIP BETWEEN TEST-TAKING STRATEGIES SELECTED AND EMOTIONAL REGULATION DIFFER ACROSS DIFFERENT LEVELS OF TEST PERFORMANCE?

Both correlations and discriminant analysis are used to address this research question. The fifteen test-taking strategies and the eight emotional regulation processes / emotions / behaviors were correlated to identify the possible relationships and compared across each of the four proficiency levels: novice, elementary, intermediate and advanced. Pearson product-moment correlation coefficients (r) were computed and Table (74) shows the significant correlations (p).

Table (74): Significant correlations between test-taking strategies and emotional regulation processes / emotions / behaviors by proficiency level

CORRELATIONS BETWEEN ...	PROFICIENCY LEVEL							
	Novice		Elem.		Inter.		Adv.	
	r	p	r	p	r	p	r	p
Dealing with unknown texts and anxiety due to time pressure	0.29	0.02	0.25	0.02				
Using L1 to deal with L2 & anxiety ...	0.28	0.02			0.31	0.01		
Using knowledge of grammar rules and anxiety due to time pressure	0.29	0.02						
Rereading text and anxiety ...	0.30	0.02						
Elimination and anxiety due to time pressure	0.32	0.01			0.25	0.02		
Making use of clues and anxiety ...			0.22	0.05				
Rereading questions and anxiety ...					0.23	0.02		
Rereading text and goal orientation	0.36	0.01					0.43	0.01
Ensuring understanding and goal orientation	0.42	0.01			0.21	0.05		
Rereading questions & reassurance provided by L1			0.23	0.05			0.33	0.05
Using L1 to deal with L2 and reassurance provided by L1					0.24	0.02		
Ensuring understanding & reassurance provided by L1			0.26	0.01				
Reading the prompt in two languages and reassurance provided by L1			0.21	0.05				

Table (74): Significant correlations between test-taking strategies and emotional regulation processes / emotions / behaviors by proficiency level (continued)

CORRELATIONS BETWEEN ...	PROFICIENCY LEVEL							
	Novice		Elem.		Inter.		Adv.	
	r	p	r	p	r	p	r	p
Managing time & reassurance provided by L1							0.43	0.01
Dealing with unknown words & reassurance provided by L1							0.37	0.02
Recognizing correct answers without hesitation & reassurance provided by L1							-0.45	0.01
Dealing with unknown texts and behaviors releasing negative emotions			0.27	0.01				
Using knowledge of grammar & behaviors releasing negative emotions					-0.22	0.05		
Using L1 to deal with L2 & assessment of confidence	0.30	0.02			-0.26	0.01		
Previewing / reviewing and assessment confidence	0.42	0.01	0.21	0.05				
Dealing with unknown texts & assessment of confidence					-0.21	0.05		
Leaving blank responses & assessment of confidence					-0.25	0.02		
Rereading questions & assessment of confidence					-0.24	0.02		
Rereading text and assessment of confidence							-0.38	0.02
Dealing with unknown texts and fear of taking tests			0.24	0.02				
Using L1 to deal with L2 and challenge	0.24	0.05			0.24	0.02		
Previewing / reviewing and challenge	0.25	0.05					0.32	0.05
Rereading text and challenge							0.33	0.05

Table (74) shows a total of 28 specific correlations between test-taking strategies and emotional regulation processes across the four proficiency levels. The picture becomes more complicated when examining the interplay between test-taking strategies, emotional regulation processes and proficiency level. It is clear that the relationship between test-taking strategies and emotional regulation does vary across proficiency

levels. None of the specific correlations were significant across the four proficiency levels and only 10 of them were significant across two of the four proficiency levels. The correlations that are specific to each proficiency level are examined.

As expected, five test-taking strategies associated with “anxiety due to time pressure” were significant for novice test-takers: “using knowledge of grammar rules”, “rereading text”, “dealing with unknown texts”, “using L1 to deal with L2” and “elimination”. “Goal orientation” was associated with the test-taking strategies of “rereading text” and “ensuring understanding”. “Assessment of confidence” was associated with “using L1 to deal with L2” and “previewing / reviewing”. Thus, test-taking strategies dealing with “anxiety due to time pressure” emerge as key relationships for novice test-takers.

For elementary level test-takers “making use of clues” and “dealing with unknown texts” was associated with anxiety. For this group of test-takers three strategies were related to “reassurance provided by L1”: “rereading questions”, “ensuring understanding” and “reading the prompt in the two languages”. “Assessment of confidence” was associated with “previewing / reviewing”. Elementary level test-takers experienced significant “fear of taking tests” and “behaviors releasing negative emotions” when handling unknown texts. The feeling of “reassurance provided by L1”, the negative emotions, anxiety and fear related to dealing with unknown texts emerge as important relationships for elementary test-takers.

The profile for intermediate level test-takers differs. “Using L1 to deal with L2” (mainly translation) was associated with several emotions and a task-focusing process: anxiety, “reassurance provided by L1” and “challenge”. Using L1 also negatively correlated with confidence. Strategies used in association with anxiety were “rereading questions” and “elimination”. “Goal orientation” correlated with the strategy of ensuring understanding. “Using knowledge of grammar rules” negatively correlated with “behaviors releasing negative emotions”, thus for intermediate test-takers, the use of grammar was a positive strategy. Four test-taking strategies correlated negatively with assessment of confidence. Thus, low confidence intermediate test-takers used the following strategies: “using L1 to deal with L2”, “dealing with unknown texts” (low confidence test-takers find difficulty in dealing with unknown texts), “leaving blank

responses” and “rereading questions”. Therefore, for intermediate test-takers using L1 emerges as a key test-taking strategy. “Using knowledge of grammar” and the feeling of high / low confidence are also significant.

Four test-taking strategies are associated with “reassurance provided by L1” for advanced level test-takers. These strategies are: “rereading questions”, “managing time”, “dealing with unknown words” and “recognizing correct answers without hesitation” (negative correlation). These relationships are difficult to interpret. However it is interesting to note that the strategy of “managing time” associated with an emotional regulation process only occurred for the advanced group. The strategy of “rereading text” was associated with “goal orientation”, “assessment of confidence” (negative correlation) and “challenge”. “Previewing / reviewing” was related to “challenge”. It was interesting to note that “using L1 to deal with L2” correlated with challenge. When advanced learners feel challenged they resort to Arabic. Therefore, for advanced level test-takers rereading emerges as a significant test-taking strategy. The feeling of reassurance provided by L1 was also significant.

Discriminant analysis was also used to further explore the relationship between test-taking strategies, emotional regulation and proficiency level.

6.8.1 Discriminant Analysis by Proficiency Level

In order to determine which test-taking strategies and emotional regulation processes discriminate between the four proficiency levels, a step-wise Discriminant Analysis (DA) was computed and the discriminating variables were the fifteen test-taking strategies and eight emotional regulation processes. The DA results are shown in Table (75). Because there are four proficiency levels, three discriminant functions are produced.

Table (75): Results of discriminant analysis by proficiency level

Discriminant Function	Eigenvalue	%Cumulative Variance	Canonical R	Wilks' Lamba	F-test	Df	P
Proficiency Level	0.83	87%	0.67	0.48	8.78	45, 1423	0.00
	0.09	97%	0.29				
	0.03	100%	0.18				

The results of the Wilks' Lamba statistic indicated that the overall test had a significant results at $p = 0.00$ level. The discriminatory power of the first two functions contributed to 97% of the variance, thus these two functions will be used in the interpretation. There is a significant difference between the group means and a significant amount of the variance (52%) accounted for by the test-taking strategies and emotional regulation processes. Other variables such as gender and age may account for the some of the variance unaccounted for.

The Jackknifed classification matrix is shown in Table (76) as follows:

Table (76): Jackknifed classification matrix for the four proficiency groups

	Novice	Elem.	Intermed.	Advanced	% Correct
Novice	46	23	3	0	64%
Elem.	62	95	50	12	43%
Intermed.	13	33	71	47	43%
Advanced	0	2	11	29	69%
Total	121	153	135	88	48%
Actual distribution	72	219	164	42	497

Because of unequal group size, it is important to examine the classifications closely. Table (76) shows that of the 72 cases in the novice group, 46 (64%) were predicted correctly to be members of the novice group and of the 42 cases in the advanced group 69% were classified correctly. For the elementary and intermediate groups only 43% were classified correctly. The overall percentage of cases classified correctly is poor: 48%.

The failure of DA to successfully distinguish between the four proficiency groups is not surprising. While it has been clearly shown from the correlations that test-takers in the different proficiency levels using significantly different test-taking strategies and emotional regulation processes, however, specific combinations of these strategies and processes cannot be used to reliably predict proficiency level (applies to only 48% of the cases). One of the main reasons is that in many cases the distinction between these four groups may be blurred. Test-takers are classified into one of these four groups based on a total score and test-takers' whose scores fall within the cut-off range between two proficiency levels may actually be wrongly placed by chance (a test-taker may be placed in a lower level by missing one item or vice versa). Therefore, it would be difficult to predict a test-taker's proficiency level on the basis of test-taking strategies and emotional regulation processes used.

Table (77) shows the canonical discriminant coefficients for each of the predictors under each function that contributed to the classification between the four groups.

Table (77): Standardized canonical discriminant coefficients for proficiency level

Discriminating Variable	Loading	
	Function 1	Function 2
Using L1 to deal with L2	0.30	0.49
Leaving blank responses	0.40	-0.04
Dealing with unknown words	0.08	0.17
Rereading text	-0.15	0.01
Handling MCQ options	-0.04	-0.42
Elimination	-0.16	0.22
Rereading questions	0.12	-0.29
Reading the prompt in two languages: L1/L2/L1	0.18	0.04
Previewing / reviewing	-0.04	0.29
Anxiety due to time pressure	0.05	0.55
Goal orientation	-0.16	0.23
Reassurance provided by L1	0.21	-0.08
Feeling relaxed	0.22	0.20
Behaviors releasing negative emotions	0.279	-0.32
Assessment of confidence	-0.57	0.18

Each of these functions discriminates between one group and the other three groups. On examining the canonical scores group means it was found that function 1 discriminates between the novice group and the other three groups while function 2 separates the elementary group. The variables that are significant i.e. loading is equal to or greater than 0.30 are examined (the absolute correlations are interpreted). The first function is related to two test-taking strategies “using L1 to deal with L2”, “leaving blank responses” and the emotional regulation process of “confidence”. From the correlations it was also found that for the novice group, using L1 correlated with assessment of confidence. Thus, translation and formulating ideas in Arabic gave confidence to these test-takers.

The second function is related to test-taking strategies “using L1 to deal with L2” and “handling MCQ options” along with two emotional regulation processes “anxiety due to time pressure” and “behaviors releasing negative emotions”. Once again this agrees with the results of the correlations which showed that elementary test-takers experienced fear, anxiety and negative emotions when “dealing with unknown texts” and “reassurance provided by L1”. Thus, translation and formulating ideas in Arabic and focusing on the options were strategies that they used in association with anxiety and behaviors dealing with negative emotions.

Therefore, both the correlations and the discriminant analysis showed that test-takers at different proficiency levels use a combination of different test-taking strategies and emotional regulation processes. However, it was not possible to predict proficiency level on the basis of a particular combination of test-taking strategies and emotional regulation processes.

The researcher had attributed the failure of DA to discriminate among the four proficiency levels to the fact that in some cases the distinction between these four groups may have been blurred. To pursue this further, the researcher decided to conduct a DA on the novice and advanced test-takers. These two groups are obviously distinct in language performance and therefore, if the assumption that specific test-taking strategies and emotional regulation processes can be used to predict test performance is true, then the DA results on the novice and advanced groups only (excluding the elementary and intermediate groups) should reflect this.

6.8.2 Discriminant Analysis by Two Proficiency Levels: Novice and Advanced
 The DA results are shown in Table (78):

Table (78): Results of discriminant analysis by two proficiency levels

Discriminant Function	Eigenvalue	Canonical R	Wilks' lamba	F-test	Df	P
Proficiency	3.60	0.89	0.22	27.69	13, 100	0.00

The results of the Wilks' Lambda statistic indicated that the overall test had a significant results at $p = 0.00$ level. The discriminatory power of the discriminatory function is expressed by the eigenvalue of 3.60 and the degree of association between the groups and the discriminant scores is expressed by the canonical correlation of 0.89. Thus, there is a significant difference between the group means and 78% of the variance has been accounted for by the test-taking strategies and emotional regulation processes / emotions / behaviors. These results are interesting and support the assumption that a combination of test-taking strategies and emotional regulation processes can be used to discriminate between novice and advanced test-takers.

The Jackknifed classification matrix is shown in Table (79) as follows:

Table (79): Jackknifed classification matrix for novice / advanced proficiency groups

	Novice	Advanced	% Correct
Novice	68	4	94%
Advanced	1	41	98%
Total	69	45	96%

Because of unequal group size, it is important to examine the classifications closely. Table (79) shows that of the 72 cases in the novice group, 68 (94%) were predicted correctly to be members of the novice group while 4 were misclassified to the advanced group. Similarly, 41 out of the 42 cases (98%) in the advanced group were

identified correctly and only 1 was assigned incorrectly to the novice group. The overall percentage of cases classified correctly is very good: 96%.

Table (80) shows the canonical discriminant coefficients for each of the predictors that contributed to the classification between the two groups.

Table (80): Standardized canonical discriminant coefficients for novice and advanced proficiency groups

Discriminating Variable	Loading
Making use of clues	0.27
Using L1 to deal with L2	0.70
Leaving blank responses	0.42
Using knowledge of grammar rules	-0.23
Rereading text	-0.55
Recognizing correct answers without hesitation	0.22
Elimination	-0.47
Ensuring understanding	-0.35
Reading the prompt in two languages: L1/L2/L1	0.35
Reassurance provided by L1	0.17
Feeling relaxed	0.24
Assessment of confidence	-0.67
Fear of taking tests	0.30

The variables with loadings of less than 0.3 are not included in the interpretation. Thus, the six test-taking strategies “using L1 to deal with L2”, “leaving blank responses”, “rereading text”, “elimination”, “ensuring understanding” and “reading the prompt in two languages: L1/L2/L1” and the emotional regulation processes / emotions of “assessment of confidence” and “fear of taking tests” are those that discriminated between the novice and advanced groups. On examining the positive and negative values of the coefficients, it can be seen that “rereading text” and “ensuring understanding” associated with “assessment of confidence” are most likely used by the advanced group while the other three strategies associated with “fear of taking tests” are used by the novice group. This DA analysis showed that it is possible to predict novice or advanced proficiency levels on the basis of a particular combination of test-taking strategies, emotional regulation processes and emotions which supports the researcher’s explanation that the distinctions between the four proficiency groups are blurred.

Therefore, both the correlations and the discriminant analyses showed a definite relationship between emotional regulation processes and test-taking strategies. Test-takers experiencing different emotions, regulation processes and behaviors do use different test-taking strategies throughout the test. The findings also show that particular combinations of test-taking strategies and emotional regulation processes, emotions and behaviors do vary by proficiency level.

The implications of the findings from the quantitative data in this chapter are discussed in section 7.1.3 in Chapter 7. The final LTP model is discussed in light of the quantitative data obtained in Chapter 5 & 6 incorporating the revisions to the model based on the findings of the think aloud data (section 5.4 discusses the fit of the LTP model to the think aloud data).

6.9 TEST-TAKING STRATEGIES, EMOTIONAL REGULATION AND TEST PERFORMANCE

Although there have been several empirical studies demonstrating the relationship between language anxiety and achievement and the use of different test-taking strategies by high and low ability test-takers, however very few studies have investigated the relationship among the test-taking strategies, affect and test performance. In one study conducted by Weir et al (2000) it was found that the higher the scores, the more interesting the text is perceived and the more frequently they used the strategies test in real life. However, this study did not focus on emotions or emotional regulation.

Similar to the findings of Hashkes & Koffman (1982), Gordon (1987), Purpura (1999) and Weir et al (2000), the findings of this study showed that high ability and low ability test-takers do use different test-taking strategies as previously discussed and which are summarized in the following section. However, the specific strategies used by high ability and low ability test-takers in the literature were not the same as identified in this study. For example, Weir et al (2000) found that readers in the top group process text very fast but this was not the case in this study. Egyptian adult test-takers at the advanced level used the strategies of translating and rereading several times.

Similar to the findings from the literature as reported by Madsen (1982), Horwitz & Young (1991) and MacIntyre et al (1997), it was found in this study that as language proficiency level increases, the occurrence of negative emotions tend to decrease.

In this study, it was found that at each of the four different proficiency levels, test-takers tended to use a combination of test-taking strategies and emotional regulation processes, emotions and behaviors. These combinations were different across the four proficiency levels. There are no studies or findings in the literature to compare with and further research is needed in this area.

6.10 SUMMARY OF FINDINGS AND OUTCOMES ON RESEARCH QUESTION 3

Based on the results and analysis above, a summary of the findings related to each sub-question of **Research Question 3** is presented below.

6.10.1 Research Question 3.1

Do test-taking strategies vary across different levels of test performance? It was found that test-taking strategies vary to some extent by proficiency level. Five of the fifteen test-taking strategies are significantly different across the four proficiency levels.

6.10.2 Research Question 3.2

If so, how do test-taking strategies differ across different levels of test performance? The advanced group use test-taking strategies of “using L1 to deal with L2” and “leaving blank answers” significantly lower than the other three groups. The novice group use the strategies of “rereading questions” and “reading the prompt first in Arabic and then in English” significantly more than the other groups. The intermediate group “used knowledge of grammar rules” most.

6.10.3 Research Question 3.3

Does emotional regulation vary across different level of test performance? Emotional regulation processes and emotions vary to a great extent by proficiency level. Five of the eight emotional regulation processes / emotions / behaviors are significantly different across the four proficiency levels.

6.10.4 Research Question 3.4

If so, how does emotional regulation differ across different levels of test performance? The advanced group experienced the highest “assessment of confidence” and the lowest “anxiety due to time pressure” and “behaviors releasing negative emotions” compared to the other groups. The novice group experienced the highest “reassurance provided by L1” and the lowest “goal orientation”.

6.10.5 Research Question 3.5

Is there a relationship between test-taking strategies selected and emotional regulation? There is some relationship between test-taking strategies used and emotional regulation. “Anxiety due to time pressure” correlates with six test-taking strategies: “making use of clues”, “dealing with unknown texts”, “using L1 to deal with L2”, “leaving blank responses”, “elimination” and “rereading questions”. “Dealing with unknown texts” correlates with three negative emotions and behaviors: “anxiety”, “behaviors releasing negative emotions” and “fear of taking tests”. “Reassurance provided by L1” correlates with the test-taking strategies of “using L1 to deal with L2”, “rereading” and “reading prompt in Arabic and then in English”. “Using knowledge of grammar rules” and “previewing / reviewing” correlate with “assessment of confidence”. “Leaving blank responses” correlates with “behaviors releasing negative emotions”.

From discriminant analysis, the four test-taking strategies: “using L1 to deal with L2”, “leaving blank responses”, “recognizing correct answers without hesitation” and “ensuring understanding” discriminate between test-takers with high and low anxiety. “Ensuring understanding” is the strategy most likely used by low anxiety test-

takers. It also means that a test-taker using this strategy is most likely experiencing low anxiety.

“Using knowledge of grammar rules”, “ensuring understanding” and “previewing / reviewing” are most likely used by high confidence test-takers. “Making use of clues”, “leaving blank responses” and “rereading questions” are likely to be used by low confidence test-takers.

6.10.6 Research Question 3.6

If so, does the relationship between test-taking strategies selected and emotional regulation differ across different levels of test performance? The relationship between test-taking strategies selected and emotional regulation differs across the four proficiency levels. From the correlations, five test-taking strategies associated with anxiety were significant for novice test-takers. For elementary test-takers, the “feeling of reassurance provided by L1” is associated with three test-taking strategies and “fear” and “behaviors releasing negative emotions” are associated with “dealing with unknown texts”. For intermediate test-takers “using L1 to deal with L2” (translation) is associated with four emotional regulation processes and emotions: “anxiety”, “reassurance provided by L1”, “assessment of confidence”(negative correlation), and “challenge”. “Assessment of confidence” correlates with four test-taking strategies. For advanced test-takers the test-taking strategy of “rereading questions” correlates with three emotional regulation processes: “goal orientation”, “reassurance provided by L1” and “assessment of confidence”. “Challenge” correlates with “previewing / reviewing” and “rereading text”.

Discriminant analysis showed that test-takers at different proficiency levels use specific combinations of test-taking strategies and emotional regulation processes. It was shown that it is not possible to predict proficiency level across four different levels from particular combinations of test-taking strategies and emotional regulation processes but that it is possible to predict novice and advanced proficiency levels.

CHAPTER 7:

CONCLUSION

This study investigated adult EFL learners' perceptions of English language proficiency, identified their test-taking strategies and emotional regulation processes during test-taking and explored the relationship between test-takers' reported use of test-taking strategies and emotional regulation on the one hand and their performance on an English language placement test on the other. The study was conducted in the Center for Adult & Continuing Education (CACE) at the American University in Cairo (AUC). In this chapter the theoretical and methodological implications of the study are discussed relating the findings to broader issues of foreign language teaching and testing within this particular context. Limitations of the study and recommendations for further research are also examined.

7.1 CONCLUSIONS FROM FINDINGS

7.1.1 Defining a Context-Specific Construct of Language Proficiency

The construct of foreign language proficiency in a specific Egyptian EFL context is defined as the ability to communicate (engage in conversations) in social settings with people from different cultural backgrounds; the ability to use English (engage in both oral and written communication) in the workplace; and the ability to understand readings obtained through written or electronic media. It was found that the three main contexts of language use of adult EFL learners are: work-related, academic and social contexts.

The findings showed that regardless of the different contexts of language use, all learners defined language proficiency in terms of fluency in speaking or in terms of some aspect of oral fluency. Fluency is the ability to speak easily without searching for words, expressing ideas quickly without thinking or hesitation as they would speak Arabic, using correct sentences. Therefore, even if a learner needs English for academic purposes, he or she would still want to speak English fluently and would want some aspect of speaking to be included in any academic English course. Teachers on the other hand reflecting the current methodology used of integrating language skills in the

classroom, defined language proficiency in terms of accuracy and fluency in the four skills for use in the workplace primarily and for academic contexts.

7.1.2 Test-Taking Strategies and Emotional Regulation in a Test-Taking Context

Most of the test-taking strategies identified in this study from the think aloud matched those found in the literature and the seventeen strategies that were identified from the think aloud data and not from the literature were mostly related to the use of L1 and were related to the specific design of the test which included directions and the writing prompt written in Arabic. There have been no studies on emotional regulation while taking language tests and thus, no comparisons of the findings of this study with the literature on language testing can be made. However, the findings of this study do support to a large extent the findings of studies conducted in education in general.

It was found that findings regarding high frequency test taking strategies differed to some extent between the think aloud and TTSQ data, however, the findings regarding low frequency test taking strategies were similar from both the think aloud and TTSQ data. The findings regarding high frequency and low frequency emotional regulation processes and emotions differed between the think aloud and TTSQ data. This may be because some test-takers were reluctant to report using specific strategies on a written questionnaire whereas during a think aloud, the process is more spontaneous and the test-taker does not have too much time to filter out or think about their responses. Another possibility is that test-takers may genuinely believe that they did not use particular strategies whereas in actual fact they did.

Using factor analysis, the TTSQ data was reduced to 15 test taking strategies and 8 emotional regulation processes, emotions and behaviors. From the qualitative data obtained from the TTSQ, five new test taking strategies and two new emotional regulation processes / emotions were identified.

From both the think aloud and TTSQ qualitative data it was found that anxiety was experienced most in the listening section followed by the reading. It was also found that the intensity of anxiety varied across sections and within one section in different items.

Based on the think aloud data, it was found that the LTP model is supported with some components revised to accurately describe the processes involved in a test taking situation. The factor analysis of the quantitative data also supported the various components of the model. The LTP model is discussed in section 7.2.4.

7.1.3 Test-Taking Strategies, Emotional Regulation and Test Performance

Five of the fifteen test taking strategies are significantly different across the four proficiency levels. This is similar to the findings from the literature which show that high ability and low ability test-takers do use different test taking strategies. However, the specific strategies used by high ability and low ability test-takers in the literature were not the same as identified in this study.

It was found that emotional regulation processes and emotions vary to a great extent by proficiency level. Five of the eight emotional regulation processes / emotions / behaviors were significantly different across the four proficiency levels. These findings are similar to the findings from the literature which report that as language proficiency level increases, the occurrence of negative emotions tend to decrease. However, in this study emotional regulation processes and emotions other than language anxiety were investigated.

In spite of the considerable research conducted in test-taking strategies and the use of different test-taking strategies by high and low ability test-takers and studies demonstrating the relationship between language anxiety and achievement, yet very few studies have investigated the relationship among the test-taking strategies, affect and language test performance.

In this study, it was found that at each of the four different proficiency levels, test-takers tended to use a combination of test-taking strategies and emotional regulation processes, emotions and behaviors. These combinations were different across the four proficiency levels. There are no studies or findings in the literature on language testing to compare with and further research is needed in this area.

7.2 IMPLICATIONS OF THE STUDY

A summary of the findings of the study related to each of the three research question is presented in sections 4.6, 5.7 and 6.9. The results presented have several implications described in this section. The aim of this study was to extend the metacognitive component of the Bachman & Palmer (1996) model by focusing on the test-taking process of adult EFL learners in a specific context.

7.2.1 Implications for Practice: Defining Context-Specific Constructs of Language Proficiency

The first step in the process is the test task. In order to ensure that the test tasks are meaningful in the particular context of this study it was important to define the construct of language proficiency underlying the test task. Prior to this study, no definition was found of language proficiency for Egyptian adult EFL learners either in the literature or within the English Studies Division at CACE, AUC. The language curriculum consisted of a series of textbooks across six proficiency stages and the placement test which was not related to the curriculum in any way consisted of a multiple-choice grammar test and a writing component. Therefore, the definition of the construct of language proficiency that emerged from this study had a significant impact on the curriculum design of the English language program and the construction of a placement test.

The language component of the Bachman & Palmer (1996) model includes four competences: grammatical, textual, illocutionary (language functions) and sociolinguistic. While the findings of this study did not extend the theoretical construct of the language component of the model, the findings and the data obtained were essential for applying this model to this particular context. From a practical perspective in order to translate these competences into an English language curriculum and an English language placement test that were relevant to the learners and met their language needs, data on defining the construct had to be collected from the parties concerned: learners and teachers in the same context. From a theoretical perspective, construct definition must be contextualized and generated from within and based on individuals'

perceptions and beliefs about language use in this particular context. Based on the data obtained from learners and teachers in the program, a construct definition of language proficiency was established. This construct is the basis of the design of the current English language curriculum and placement test at CACE, AUC. The construct is presented in Chapter 4, section 4.5.

7.2.2 The Relationship between Test-taking Strategies and Emotional Regulation

Fifteen test-taking strategies and eight emotional regulation processes, emotions and behaviors were identified by factor analysis. This study showed that ten of the fifteen test-taking strategies are associated with specific emotional regulation processes and emotions. For example, anxiety is associated with six test-taking strategies examples of which are leaving blank responses and rereading. “Assessment of confidence” is associated with the test-taking strategies of “using knowledge of grammar rules” and “previewing / reviewing”. In other words, test-takers who experience specific emotions either positive or negative, tend to select a specific combination of test-taking strategies. The research findings illustrate that particular test-taking strategies are not necessarily exclusive to either positive or negative emotions. In fact, some test-taking strategies are associated with both positive and negative emotions. For example, the strategies of rereading questions and translation (using L1 to deal with L2) are associated with both anxiety (a negative emotion) and “reassurance provided by L1” (a process associated with a positive emotion). Therefore, it is difficult to establish the relationship between particular test-taking strategies and either positive or negative emotions. Thus, “using L1 to deal with L2” associated with “assessment of confidence” will lead to different results on a test as compared with using the same strategy associated with anxiety.

The discriminant analysis findings also showed that test-takers who experienced different intensities of emotions i.e. they were either at the high or low end of a particular emotion / process continuum, used different test-taking strategies. It was shown that test-takers who experienced high anxiety or high confidence used different test-taking strategies as compared with those who were in the low anxiety or low assessment of confidence group.

It is therefore possible to conclude that emotional regulation influences cognitive processes i.e. the selection of particular test-taking strategies, which in turn is associated with performance on a language placement test. This means that test-takers who experience high anxiety for example, tend to use particular test-taking strategies and this correlated with the level of language proficiency of these test-takers.

7.2.3. The Relationship between Test-taking Strategies, Emotional Regulation and Test Performance

One of the most important findings in this study is the relationship between test-taking strategies, emotional regulation and test performance. It has been shown that novice, elementary, intermediate and advanced level test-takers use different combinations of test-taking strategies in association with specific emotional regulation processes and emotions. For example, it was found that for novice test-takers anxiety was associated with five test-taking strategies; self-satisfaction associated with two test-taking strategies and confidence with two test-taking strategies. For advanced level test-takers the feeling of relief was associated with four test-taking strategies; the strategy of rereading text was associated with three emotional regulation processes / emotions and the strategy of previewing / reviewing was associated with challenge. Novice and advanced test-takers used the same test-taking strategies but not associated with the same emotional regulation processes. For example, novice test-takers used the strategy of rereading associated with anxiety resulting in poor performance while advanced test-takers used rereading associated with confidence resulting in high test performance. Therefore, within this specific Egyptian EFL context, it is the particular combinations of test-taking strategies and emotional regulation processes that can be used to predict proficiency level for this particular group of test-takers.

Discriminant analysis (DA) was used to examine whether test-taking strategies, emotional regulation processes, emotions and behaviors were predictors of proficiency level at the four different stages: novice, elementary, intermediate and advanced. The DA findings showed that this was not possible. Because the researcher attributed the failure of DA to discriminate among the four proficiency levels to the fact that in some cases the distinction between these four groups may have been blurred, another DA was

conducted on test-takers at the two ends of the proficiency continuum: novice and advanced groups. Findings showed that “confidence” associated with “rereading text” and “ensuring understanding” distinguished the advanced group of test-takers from the novice group. On the other hand, the novice group is characterized by “fear of taking tests” associated with “using L1 to deal with L2”, “leaving blank responses”, “elimination” and “reading the prompt in two languages: L1/L2/L1”.

What is clear is that novice test-takers experience far more negative emotions compared to advanced test-takers, however, it is difficult to ascertain whether it is the negative emotions that cause low test performance or whether low ability test-takers will always experience negative emotions.

Another important conclusion is that different specific combinations of test-taking strategies and emotional regulation processes are obtained when examining the relationship between test-taking strategies and emotional regulation, compared to the results of examining the interrelationships among the three dimensions of the test-taking situation: strategies, emotional regulation and test performance. For example, when only examining the relationship between test-taking strategies and emotional regulation, it was found that high anxiety is associated with a test-taking strategy such as “using L1 to deal with L2”. When adding the third dimension of language proficiency, it was found that “using L1 to deal with L2” was associated with anxiety for novice and intermediate level test-takers. It was also found that this strategy was associated with challenge for the advanced level test-takers. Therefore, examining the test-taking situation 3-dimensionally (test-taking strategies, emotional regulation and proficiency level) leads to a far richer depiction of the complex nature of the interrelationships and results in a better understanding of the processes involved.

7.2.4 The Language Testing Processing (LTP) Model

Another significant outcome of this study is that a contribution has been made by extending the metacognitive strategies component of the Bachman & Palmer (1996) model. The Bachman & Palmer model illustrates the interaction between characteristics of individuals which include language knowledge, topical knowledge, personal characteristics and metacognitive strategies, and characteristics of the test task. The

metacognitive strategies consist of goals setting, assessment and planning. The model also includes affective schemata which consist of an individual's emotions. However, the model does not include cognitive test-taking strategies.

The LTP model takes an expanded view of the language test-taking process to include self-regulation of cognitive test-taking strategies and emotions, test-taking behavior and consequences of test performance. This is the first published study (to the best of the researcher's knowledge) in language testing where emotional regulation is included as a component in the test-taking process. The LTP model was initially postulated based on the literature incorporating the following elements from the Bachman & Palmer model: test task, goals, language knowledge, topical knowledge, personal characteristics and the metacognitive strategies component relabeled as assessment; and elements from educational psychology specifically: self-regulation and emotions during test-taking. The model was then empirically validated based on think aloud protocols. The relationships between test-taking strategies, emotional regulation and proficiency were further validated by quantitative data as previously discussed. The factor analysis of the quantitative data defined the construct of emotional regulation as consisting of two negative emotions, two positive emotions, one task-focusing process and two assessment processes. The behavioral component within the model was also identified as a factor in the data. The quantitative data analysis showed that emotional regulation does affect the selection of test-taking strategies used and this is associated with performance on a test.

Schutz & Davis (2000:253) had posed several questions related to coping with pleasant emotions: "Do we regulate to keep goal-congruent emotions going?... How do we regulate to keep from getting too happy and overconfident during a test?" The findings in this study from both the think aloud and quantitative data showed that when the test is going well for a test-taker they experience positive emotions and they use specific task-focusing processes and test-taking strategies to complete the test task at hand. The findings did not show how test-takers regulate overconfidence and feeling too happy or whether they actually experienced these emotions. Some test-takers noted that they did not experience any anxiety while taking the test while others (seventeen respondents) stated that they did not "feel anything in particular, they just took the test."

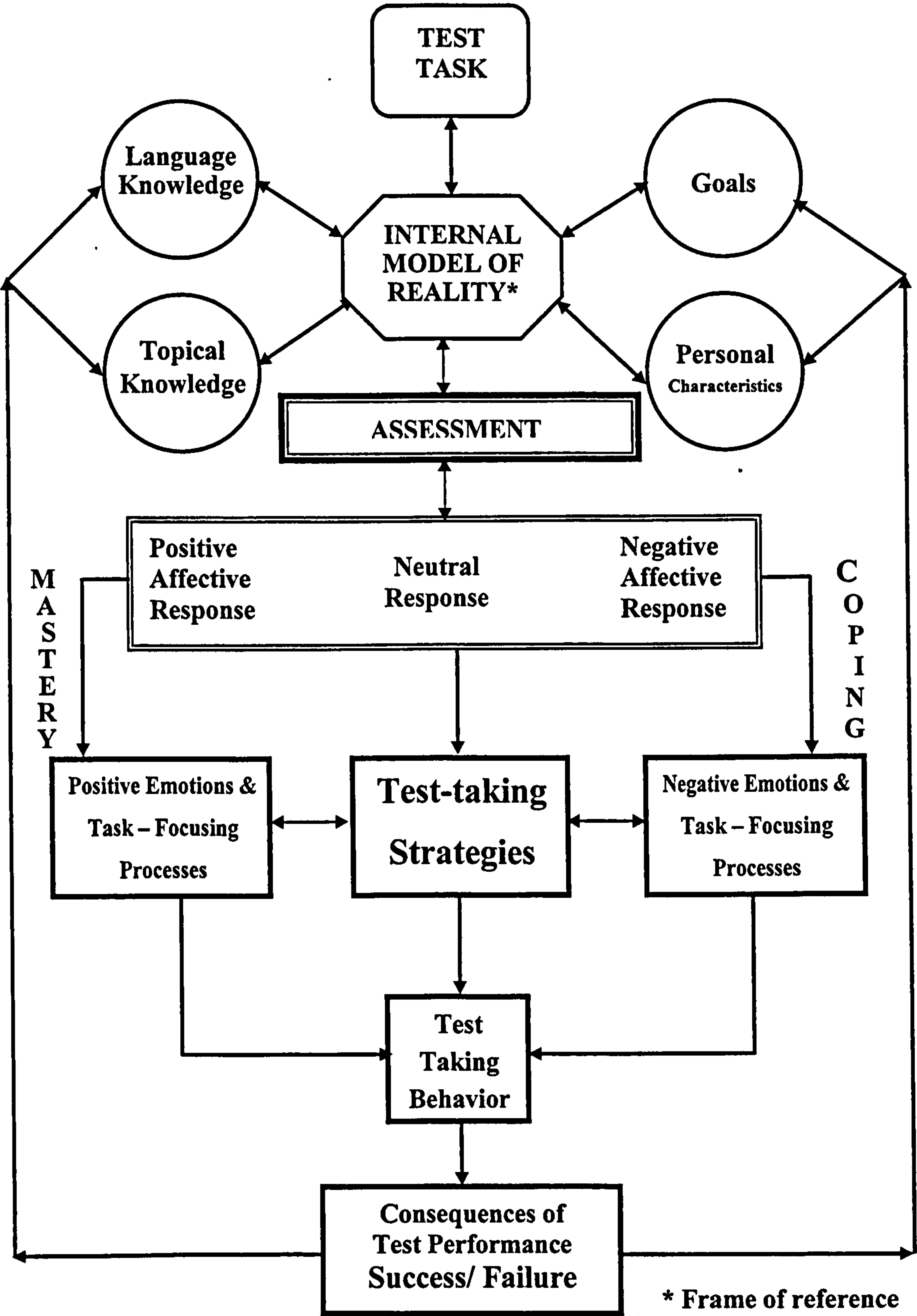
Figure (7) illustrates the final LTP model taking into account the findings of the think aloud and quantitative data. The model illustrates the process for one test task. The model is cyclical where the process is repeated for each test section and test item throughout a test. The test-taker starts out by analyzing the requirements of the test task and evaluates his / her ability to complete the task. This assessment is based on topical knowledge and personal characteristics and parallel to this process, the test-taker also makes comparisons between his / her goals and the test situation. The factor analysis of the data identified a goal orientation and an assessment process. The results of these assessments may or may not lead to a discrepancy between the requirements of the test task and language knowledge, topical knowledge and personal characteristics. If there is no discrepancy and the test is judged as not important to the test-taker's goals, there is no positive or negative affective response. From the qualitative questionnaire data, several test-takers noted a neutral response. If the test is judged as being an important and relevant goal, it may lead to a positive or negative affective response.

If there is no discrepancy and the test is judged as important and relevant, then there is a positive affective response and this is labeled the mastery mode where the test take is in control. Positive emotions are experienced and the test-taker uses test-taking strategies and task-focusing processes on the test tasks. The quantitative data analysis showed the different strategies, emotions and task-focusing processes used by test-takers at different levels of performance.

If there is a discrepancy and the test is seen as an important goal, a negative affective response occurs and this is labeled the coping mode where negative emotions emerge and task-focusing processes are used.

In both modes the test-taker selects a variety of test-taking strategies to be used (there is no differentiation between contributory and non-contributory strategies). The interplay between the test-taking strategies selected, the emotions that emerge and the task-focusing processes used determines the behavior of the test-taker. Possible behaviors are: increasing / decreasing effort, persisting in completing the task, giving up or seeking help (asking for answers or even attempting to cheat during the test). The consequences of test performance are either successful or unsuccessful experiences which feedback once again into the sources of information used for assessing capability.

Figure (7): The final language testing processing (LTP) model
(after Bachman & Palmer, 1996)



Emotional regulation during test-taking is a sequence of activities or processes that involve emotions and cognitive strategies used, designed to attain or achieve success on a test / test task. This interaction between test-taking and emotional regulation has been shown to be associated with test performance.

7.2.5 Learner Awareness of Test-taking Strategies and Emotional Regulation

Learning a foreign language is a complex process. Research in EFL teaching methods has shown that consciousness-raising classroom activities have enhanced language learning and performance (Schmidt, 1990; Fotos, 1993). It has been found that explicit training in language learning strategies improved performance and many teachers now incorporate learning strategies in their lessons whenever appropriate (Purpura, 1999; Oxford, 1990). Tests have important consequences for learners and in recent years the focus on tests has increased and is still growing. Therefore, understanding the processes that are involved in a test-taking situation, especially an English language placement test, is very important for teachers and learners. Given the positive results of language learning strategy training, by extension, test-taking strategy training would be equally effective for the learner. Tests by their very nature are stressful encounters. Accordingly, such encounters are associated with both positive and negative emotions. This study has shown that emotional regulation interacts with test-taking strategies and this interaction in turn, is associated with test performance. Therefore, it can be concluded that in the context of test-taking, training in test-taking strategies and emotional regulation would enhance test performance. If the test performance of a test-taker suffers in any way from the test-taker's inability to regulate emotions or use effective test-taking strategies, then strategy and emotion regulation training will help students approach a test effectively. The test results would then better reflect the test-taker's actual ability.

In this study, test-taking strategies and emotional regulation are not conceptualized as a list of "effective" or "ineffective" behaviors, however, they are presented within a larger and more complex system of inter-related processes and behaviors. Therefore, when thinking about test performance teachers should focus on

learners' awareness of both their cognitive and metacognitive strategies as well as emotions. Involving learners in awareness raising activities will encourage them to take responsibility for their own learning and provide them with a greater degree of autonomy.

7.2.6 Test-taking Strategy and Emotional Regulation Training

In view of the findings from the study, test-taking strategy and emotional regulation training should be an important component of teacher preparation courses. The first step in implementing strategy and emotional regulation training is for teachers to explain to learners what is meant by test-taking strategies and emotional regulation and they should stress the significant of using strategies and emotional regulation in achieving their goals. Teachers should elicit the strategies and emotional regulation processes their learners are currently using through a number of methods: class discussions, think aloud and questionnaires. It is important for teachers to sensitize learners to the range of different test-taking strategies they can use and to raise their awareness regarding the emotions that emerge during a test and processes for handling these emotions. Teachers should encourage learners to focus or "notice" their own performance and to help them develop ways to talk about test-taking strategies and emotional regulation processes. This is especially important in the Egyptian education context where education reform efforts focusing on student autonomy are currently being implemented.

Therefore, effective teacher preparation courses should include a component on developing their learners' self-awareness of test-taking processes. This should also include helping learners to self-evaluate. Here self-assessment differs from the traditional conceptualization of self-assessment, where learners are asked to evaluate their language ability in specific skills or content area. Rather, self-assessment here would mean that the teacher asks learners to focus on whether their answer was correct or not and to analyze the strategies and processes used while completing the task.

The value of this study is that it provides specific information that EFL teachers preparing Egyptian adults to sit for an English language proficiency test can use when guiding students in developing effective strategies and emotional regulation processes.

The results of this study illustrate the extent to which the combination of strategy use and specific emotional regulation processes / emotions / behaviors may have a differential effect on test performance.

7.2.7 Implications for Test Developers

Applying a framework of language ability at the stage of designing a test is not always easy sometimes because of the complexity of the framework and because of the difficulty of isolating and testing specific language features. Language performance is actually a dynamic interplay of several language features. One of the implications of this study is that when designing tests, test developers should also focus on the test process itself and not just on the design features or the psychometric properties of test items. For example, when designing multiple-choice tests, typically items in the test are organized in order of difficulty with the easier items appearing first followed by the more difficult items. The findings of this study showed that perceptions of item difficulty which is related to the different degrees of anxiety experienced, varied throughout the test. Test-takers' feelings of anxiety varied from section to section and in some cases varied within one section. From the think-aloud data, one test-taker found the last few items in the listening section to be easier than the first items. Thus, on examining the test-taking process itself, the results show that there is no need to organize items based on item difficulty since individual test-takers feelings and reactions to different items vary considerably. Another example where the test process may help test developers in organizing the various sections of a test is to examine test-takers' feelings and reactions to the overall sections. If a specific group of test-takers experience anxiety most in the reading section, then perhaps the test developer may decide to place the reading section at the beginning of the test rather than at the end because fatigue confounded by high anxiety will probably adversely affect performance. Test-takers' performance may be enhanced by placing the high anxiety section at the beginning of the test.

7.2.8 Methodological Implications

This study illustrated the value of combining both quantitative and qualitative approaches of research. Both approaches were integrated in most phases of the study. Relying solely on one or the other of the approaches would have been limiting, especially when investigating emotions in test-taking which is a difficult area to explore and quantify. The insights gained from the think aloud data enabled the researcher to interpret some of the factors identified in the factor analysis that were difficult to interpret. For example, Factor 15 of the test-taking strategies was a challenge to interpret. The two items that loaded on it did not seem to be related: having sufficient time to read the options of the next question before listening to the question in the listening task and rewriting words or phrases in the writing task. Having analyzed the think aloud data and gained an understanding of the underlying test-taking processes, the researcher was able to interpret and label the factor. Furthermore, both the qualitative and quantitative data obtained were used to validate the LTP model and this would not have been possible by relying on only one approach.

This study demonstrated the value of discriminant analysis for investigating the relationships between test-taking strategies, emotional regulation and test performance. Discriminant analysis proved to be an effective tool in classifying test-takers along a continuum of high or low emotion / process in terms of the test-taking strategies used. Discriminant analysis was used to classify test-takers along a continuum of high or low language proficiency in terms of specific combinations of test-taking strategies and emotional regulation processes / emotions / behaviors. Exploratory factor analysis was also an effective tool used to reduce the items on the TTSQ by identifying clusters and creating composite variables which could then be explored further with discriminant analysis.

7.3 LIMITATIONS OF THE STUDY

It is difficult to capture all aspects of the processual and dynamic nature of emotional regulation and therefore, there may be aspects or components of emotional regulation during test-taking that have not been elicited or included in the data collected. Traditional teaching methods are still widespread in the Egyptian educational system

and therefore the skills of self-assessment and awareness of learning and test-taking strategies are not well developed among many Egyptians. Therefore, test-takers lacking strategy training and self-regulation skills may not have been able to articulate explicitly the full range of test-taking strategies used and emotional regulation experiences. Accordingly, the LTP model may not be comprehensive and potentially influential processes, strategies or behaviors may have not been included. However, this study does not claim to be exhaustive nor comprehensive and the findings supported the LTP model given the strategies / processes / emotions / behaviors that were identified and measured.

The effects of emotions, task-focusing processes and assessment processes impacting upon one another are difficult to disentangle at times in the data as they relate to test-taking strategies. For example, it was difficult to interpret the finding that the following four test-taking strategies “rereading questions”, “managing time”, “dealing with unknown words” and “recognizing correct answers without hesitation” (negative correlation) are associated with “reassurance provided by L1” for advanced level test-takers.

It is inevitable in research in this area that cause and effect is difficult to establish. Thus, a question for future research remains: does the association of particular test-taking strategies and emotional regulation processes with lower distress and higher test performance mean that emotional regulation reduces distress or do test-takers with a particular emotional regulation system tend to use particular strategies that leads to high test performance?

All the respondents of the study were Egyptian adult EFL learners at the Center for Adult Continuing Education of the American University in Cairo. Therefore, the results of this study can only be relatable or transferable to similar contexts where the native language is Arabic and where English is a foreign language.

7.4 RECOMMENDATIONS FOR FURTHER RESEARCH

More research is needed in refining the construct of emotional regulation in order to identify further components. The application of research methods such as structural equation modeling in addition to carrying out more think aloud studies that include

videotaping would be useful in further exploring emotional regulation during test-taking. Videotaping may capture more behaviors and emotional regulation processes than audiotaping.

A second area in need of research is investigating the impact of test-takers' characteristics such as gender, age, educational background and previous experience with tests on the assessment processes that occur during a test. The effect of these variables on emotional regulation should be investigated especially in this context where gender may be a key variable.

Because the LTP model was based on several theoretical frameworks and the findings appeared to support the model within the specific Egyptian EFL context, an additional area of further research would be to investigate the applicability of the model to other test contexts outside of Egypt. The occurrence of emotions in different cultural settings could be explored and compared. The LTP model should also be researched using a variety of language test tools beyond multiple-choice or a short writing test. Thus, a taxonomy of the structure of emotions related to the structure of tasks could be established. The LTP model could also be applied to standardized tests to examine the impact on students academically and emotionally. Another area of research where a specific feature of the LTP model could be investigated is the effect of goal orientation, the perceived consequence of tests to test-takers (what the test means to the test-taker in terms of results), on test-taking strategies and emotional regulation. The relationship between test consequence and test-taking strategies and emotional regulation have implications on classroom practice and technical issues in test design.

The LTP model does not claim to be comprehensive and only provides a description of specific constructs within the complex process of taking a test. More theories are needed to drive research and better understand the relationship between emotion, cognition and other theoretical constructs such as motivation in a test-taking context.

Another area of research that is important is to investigate the effectiveness of test-taking strategy and emotional regulation training on test performance and whether enhancing learners' awareness of the value of using test-taking strategies and regulating

their emotions has an impact on their performance. Research is needed on finding out the effectiveness of different designs of training interventions.

7.5 SUMMARY OF MAJOR FINDINGS

The key findings from this study are summarized as follows:

- The definition of the construct of language proficiency at CACE, AUC was established based on learners' and teachers' perceptions and beliefs about language use in this particular context. This construct is the basis of the design of the current English language curriculum and placement test at CACE, AUC.
- The test-taking strategies, emotional regulation processes, emotions and behaviors used by adult EFL learners in a particular context while taking an English language placement test were identified.
- Emotional regulation influences the selection of particular test-taking strategies.
- Test-takers at different language proficiency levels use different combinations of test-taking strategies in association with specific emotional regulation processes and emotions.
- A Language Testing Processing (LTP) model describing the process of taking a test in terms of test-taking strategy use and emotional regulation was refined and supported based on both qualitative and quantitative data.

REFERENCES

- Aida, Y. (1994) Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. The Modern Language Journal, 78, pp. 155-168
- Alderson, C. (1991) Language testing in the 1990s: How far have we come? How much further have we to go? In: Anivan, S. (ed.) Current Developments in Language Testing. Singapore: Regional Language Center, pp.1-26
- Alderson, C. (1990) Testing reading comprehension skills (Part Two): Getting students to talk about taking a reading test (a pilot study). Reading in a Foreign Language 7(1), pp. 465-503
- Alderson, C. and Clapham, C. (1992) Applied linguistics and language testing: A case study of the ELTS test. Applied Linguistics, 13(2), pp. 149-167
- Allan, A. (1995) Begging the questionnaire: instrument effect on readers' responses to a self-report checklist. Language Testing, 12, pp. 133-156
- Allan, J. (1996) Curriculum design in higher education using a learning outcome-led model: its influence on how students perceive learning. Unpublished Ph.D. thesis, University of Wolverhampton, England
- Allen, P., Cummins, J., Mougeon, R. and Swain, M. (1983) The Development Of Bilingual Proficiency: Second Year Report And Appendices. Toronto: Modern Language Center OISE
- Anastasi, A. (1990). Ability testing in the 1980's and beyond: Some major trends. Public Personnel Management, 18(4),pp.471-485
- Anastasi, A. (1988) Psychological Testing. 6th ed., New York: Macmillan
- Anastasi, A. (1986) Evolving concepts of test validation. Annual Review of Psychology, 37, pp.1-15
- Anderson, N., Bachman, L., Perkins, K. and Cohen, A. (1991) An exploratory study into the construct validity of a reading comprehension test: Triangulation of data sources. Language Testing, 8, pp.41-66
- Anderson, P.L. (1982) Self-esteem in the foreign language: A preliminary investigation. Foreign Language Annals, 15, 109-115
- Aoki, N. (1999) The role of affect in the development of learner autonomy. In: Arnold, J.(ed.) Affect in Language Learning. New York: Cambridge
- Aydelott, J., El Bakary, W., Boraie, D., Calderbank, R., Laurence, M. and Said, I. (2000) Reconstructing an EFL curriculum from ground up. Paper presented at the TESOL 34th Annual Convention, Vancouver, B.C., Canada
- Bachman, L. (2000) Modern language testing at the turn of the century: assuring that what we count counts. Language Testing, 17(1), pp.1-42
- Bachman, L. (1990) Fundamental Considerations in Language Testing. Oxford: Oxford University Press

- Bachman, L. (1988) Problems in examining the validity of the ACTFL oral proficiency interview. Studies in Second Language Acquisition, 10, pp.149-164
- Bachman, L and Cohen, A. (1998) Language testing – SLA interfaces: An update. In: Bachman, L. and Cohen, A. (eds.) Interfaces between Second Language Acquisition and Language Testing Research. Cambridge: Cambridge University Press, pp.1-31
- Bachman, L. and Palmer, A. (1996) Language Testing in Practice. Oxford: Oxford University Press
- Bachman, L. and Palmer, A. (1989) The construct validation of self-ratings of communicative language ability. Language Testing, 6, pp.14-25
- Bachman, L. and Palmer, A. (1981) The construct validity of the FSI Oral Proficiency Interview. Language Learning, 31, pp.67-86
- Bachman, L. and Savignon, S. (1986) The evaluation of communicative language proficiency: a critique of the ACTFL oral interview. Modern Language Journal, 70, pp. 380-390
- Banerjee, J. and Luoma, S. (1997) Qualitative approaches to test validation. In: Clapham, C. and Corson, D. (eds.) Encyclopedia of Language and Education, Volume 7: Language Testing and Assessment, pp. 275-287
- Bassey, M. (1990) On the nature of research in education (Part 2). Research Intelligence, 37, Summer
- Blanche, P. (1990) Using standardized achievement and oral proficiency tests for self-assessment purposes: the DLIFLC study. Language Testing, 7, pp.202-209
- Blanche, P. (1986) The relationship between self-assessments and other measures of proficiency in the case of adult foreign language learners. Unpublished master's thesis, University of California, Davis, California
- Blanche, P. and Merino, B. (1989) Self-assessment of foreign language skills: implications for teachers and researchers. Language Learning, 39, pp.313-340
- Blue, G.M. (1988) Self-assessment: The limits of learner independence. ELT Documents 131, 100-118
- Boekaerts, M. (1991) Subjective competence, appraisals and self-assessment. Learning and Instruction, Vol. 1, pp.1-17
- Boekaerts, M. and Niemivirta, M. (2000) Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of Self-Regulation. San Diego, CA: Academic, pp.417-450
- Boraie, D., Kassabgy, O. and Schmidt, R. (1993) Motivation in EFL: The case in Egypt. Paper presented at the TESOL 27th Annual Convention, Atlanta, Georgia
- Bradshaw, J. (1990) Test takers' reactions to a placement test. Language Testing, 7, pp.13-30
- Brindley, G. (ed.) (1995) Language Assessment in Action. Research Series 8, NCELTR, Macquarie University, Sydney NSW 2109

- Brindley, G. (1991) Defining language ability: the criteria for criteria. In: Anivan, S. (ed.) Current Developments in Language Testing. Singapore: Regional Language Center, pp.139-164
- Brown, A. (1993) The role of test taker feedback in the test development process: Test takers' reactions to a tape-mediated test of proficiency in spoken Japanese. Language Testing, 10, pp.277-304
- Buck, G. (1994) The appropriacy of psychometric measurement models for testing second language listening comprehension. Language Testing, 11(2), pp.145-170
- Buck, G. (1992) Translation as a language testing procedure: Does it work? Language Testing, 9, pp. 123-148
- Byrne, D. (2002) Interpreting Quantitative Data. London: Sage
- Canale, M. (1983) On some dimensions of language proficiency. In: Oller, J. (ed.) Issues in Language Testing Research. Rowley, MA: Newbury House, pp.333-342
- Canale, M. and Swain, M. (1980) Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1, pp.1-47
- Carroll, J. (1983) Psychometric theory and language testing. In: Oller, J. (ed.) Issues in Language Testing Research. Rowley, MA: Newbury House
- Carroll, J. (1961) Fundamental considerations in testing for English language proficiency of foreign students. Testing the English Proficiency of Foreign Students. Washington, DC: Center for Applied Linguistics, pp. 30-40
- Carver, C.S. and Scheier, M.F. (2000) On the structure of behavioral self-regulation. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of Self-Regulation. San Diego, CA: Academic, pp.42-85
- Chalhoub-Deville, M. (1997) Theoretical models, assessment frameworks and test construction. Language Testing, 14(1), pp. 3-22
- Chalhoub-Deville, M. (1995) Deriving oral assessment scales across different tests and rater groups. Language Testing, 12, pp.16-33
- Clapham, C. (1981) Reaction to the Carroll paper (1). In: Alderson, C. and Hughes, A.(eds.) Issues in Language Testing. ELT Documents 111. London: British Council
- Clement, R. and Kruideneir, B.G. (1985) Aptitude, attitude and motivation in second language proficiency: A test of Clement's model. Journal of Language and Social Psychology, 4, pp.21-37
- Cohen, A. (2000) Exploring strategies in test-taking: Fine-tuning verbal reports from respondents. In: Ekbatani, G. and Pierson, H. (eds.) Learner-Directed Assessment in ESL. Mahwah, New Jersey: Lawrence Erlbaum Associates, pp. 127-164
- Cohen, A. (1998) Strategies and processes in test taking. In: Bachman, L. and Cohen, A. (eds.) Interfaces between Second Language Acquisition and Language Testing Research. Cambridge: Cambridge University Press, pp. 90-111

- Cohen, A. (1994) English for academic purposes in Brazil: The use of summary tasks. In: Hill, C. and Parry, K. (eds.) From Testing to Assessment: English as an International Language. London: Longman, pp. 174-204
- Cohen, A. (1984) On taking language tests: What the students report. Language Testing, 1, pp.70-81
- Cohen, A. and Aphek, E. (1979) Easifying second language learning. Research report under the auspices of Brandeis University and submitted to the Jacob Hiatt Institute, Jerusalem. ERIC Document ED 163753
- Cohen, L. and Manion, L. (1994) Research Methods in Education. 4th ed. London: Routledge
- Coombe, C.A. (1992) The relationship between self-assessment ratings of functional literacy skills and basic English skills test results in adult refugee ESL learners. Ph.D. Thesis, Ohio State University. Dissertation Abstracts International, 53, 3774A
- Council of Europe. (2001) Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge: Cambridge University Press
- Creswell, J.W. (1994) Research Design: Qualitative and Quantitative Approaches. Thousand Oaks: Sage
- Creswell, J.W. and Miller, D.L. (2000) Determining validity in qualitative inquiry. Theory into Practice, 39(3), pp.124-130
- Cronbach, L.J. (1989) Construct validation after thirty years. In: Linn, R. (ed.) Intelligence: Measurement, Theory and Public Policy. Urbana: University of Illinois Press, pp.147-171
- Csikszentmihalyi, M. and Nakamura, J. (1989) The dynamics of intrinsic motivation: a study of adolescents. In: Ames, C. and Ames, R. (eds.) Research on Motivation in Education, (Vol. 3: Goals and cognitions). San Diego: Academic Press
- Cumming, A. (1996) Introduction: The concept of validation in language testing. In: Cumming, A. and Berwick, R. (eds.) Validation in Language Testing. Clevedon, UK: Multilingual Matters, pp.1-14
- Dandonoli, P. and Henning, G. (1990) An investigation of the construct validity of the ACTFL proficiency guidelines and oral interview procedure. Foreign Language Annals, 23, pp.11-22
- Davidson, F. and Henning, F. (1985) A self-rating scale of English difficulty: Rasch scalar analysis of items and rating categories. Language Testing, 2, pp.164-191
- Davies, A. (1996) Outing the tester: Theoretical models and practical endeavours in language testing. In: Blue, G.M. and Mitchell, R. (eds.) Language and Education. Clevedon, Avon: British Association for Applied Linguistics in association with Multilingual Matters, pp.60-69
- Davies, A. (1981) Review of J. Munby: Communicative Syllabus Design. TESOL Quarterly, 15(3), pp. 332-336
- Denzin, N.K. (1978) The Research Act: A Theoretical Introduction to Sociological Methods. New York: McGraw-Hill

- Dickinson, L. (1995) Autonomy and motivation: A literature review. System, 23, pp.165-174
- Douglas, D. (2001) Performance consistency in second language acquisition and language testing research: a conceptual gap. Second Language Research, 17(4), pp. 442-456
- Ehrman, M. (1998) Affect, cognition, and learner self-regulation in second language learning. In: Kagan, O. and Rifkin, B.(eds.) The Learning and Teaching of Slavic Languages and Cultures: Toward the 21st Century. Bloomington: Slavica, pp.1-15
- Ehrman, M. (1996) Understanding Second Language Learning Difficulties: Looking Beneath the Surface. Thousand Oaks, CA: Sage
- Emanuel, E. (1982) The use of the cloze technique – regular cloze and discourse cloze – in the teaching of Hebrew as a mother tongue. Course paper (in Hebrew), School of Education, Hebrew University of Jerusalem, Jerusalem, Israel
- Endler, N.S. and Kocovski, N.L. (2000) Self-regulation and distress in clinical psychology. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of Self-Regulation. San Diego, CA: Academic, pp.569-600
- Ericsson, K.A. and Simon, H.A. (1993) Protocol Analysis: Verbal Reports as Data. Cambridge, MA: MIT Press
- Flick, U. (2002) An Introduction to Qualitative Research. London: Sage
- Fotos, S. (1993) Consciousness-raising and noticing through focus on form: grammar task performance versus formal instruction. Applied Linguistics, 14, pp.385-407.
- Gagne, E.D., Yekovich, C.W. and Yekovich, F.R. (1993) The Cognitive Psychology of School Learning. New York, NY: HarperCollins College Publishers
- Gall, M., Borg, W. & Gall, J. (1996). 6th edition. Educational Research An Introduction. New York: Longman Publishers USA
- Gardner, R.C., Moorcroft, R. and Metford, J. (1989) Second language learning in an immersion programme: Factors influencing acquisition and retention. Journal of Language and Social Psychology, 8, 287-305
- Gordon, C. (1987) The effect of testing method on achievement in reading comprehension tests in English as a foreign language. Master's thesis, School of Education, Tel Aviv University, Tel Aviv, Israel
- Green, A. (1998) Verbal protocol analysis in language testing research: A handbook. In: Milanovic, M. (ed.) Studies in Language Testing, 5. Cambridge: University of Cambridge Local Examinations Syndicate and Cambridge University Press
- Guba, E. and Lincoln, Y. (1994) Competing paradigms in qualitative research. In: Denzin, N. and Lincoln, Y. (eds.) Handbook of Qualitative Research. London: Sage
- Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, R.C. (1998) Multivariate Data Analysis, (5th ed.) Upper Saddle River, NJ: Prentice-Hall
- Hamp-Lyons, L. and Lynch, B.K. (1998) Perspectives on validity: a historical analysis of language testing conference abstracts. In: Kunnan, A.J. (ed.) Validation in Language Assessment. New Jersey: Lawrence Erlbaum Associates

Harley, B., Allen, P., Cummins, J. and Swain, M. (1990) (eds.) The Development of Second Language Proficiency. New York: Cambridge University Press

Hashkes, B. and Koffman, N. (1982) Strategies used in a cloze test. Course paper, School of Education, Hebrew University of Jerusalem, Jerusalem, Israel

Heilenmann, L. (1991) Self-assessment and placement: A review of the issues. In Teschner, R. (ed.) Assessing Foreign Language Proficiency of Undergraduates. Boston: Heinle and Heinle, pp.93-114

Heilenmann, L. (1990) Self-assessment of second language ability: The role of response effects. Language Testing, 7, pp.174-201

Hill, C. and Parry, K. (1994) Models of literacy: the nature of reading tests. In; Hill, C. and Parry, K. (eds.) From Testing to Assessment: English as an International Language. London: Longman Group Ltd

Hinofotis, F., Bailey, K., and Stern, S. (1981) Section II. Empirical research. In: Palmer, A., Groot, P. and Tropper, G. (eds.) The Construct Validation of Tests of Communicative Competence. Washington, DC, pp.106-126

Horwitz, E.K. and Young, D.J. (eds.) (1991) Language Anxiety: From Theory and Research to Classroom Implications. Englewood Cliffs, NJ: Prentice-Hall

Howe, K. (1998) The interpretive turn and the new debate in education. Educational Researcher, 27(8), pp. 13-20

Hughes, A. (1989) Testing for Language Teachers. Cambridge, UK: Cambridge University Press

Israel, A. (1982) The effect of guessing in multiple-choice language tests. Course paper, School of Education, Hebrew University of Jerusalem, Jerusalem, Israel

Janssen-van Dieten, A.M. (1992) Zelfbeoordeling en tweede-taalen (Self-assessment in second language learning). Unpublished Ph.D. thesis, Katholieke Universiteit, Nijmegen, The Netherlands

Janssen-van Dieten, A.M. (1989) The development of a test of Dutch as a second language: the validity of self-assessment by inexperienced subjects. Language Testing, 6, pp.1-13

de Jong, J. and Verhourven, L. (1992) Modeling and assessing language proficiency. In: Verhoeven, L. and de Jong, J. (eds.) The Construct of Language Proficiency. Amsterdam: John Benjamin Publishing Company

de Jong, J. (1987) Defining tests for listening comprehension: a response to Dan Douglas's 'Testing listening comprehension'. In: Valdman, A. (ed.) Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency. Bloomington, IN: Indiana University

Kenyon, D. (1995) An investigation of the validity of the demands of tasks on performance-based tests of oral proficiency. Paper presented at the meeting of the Language Testing Research Colloquium, Long Beach, CA

- Kenyon, D. and Stansfield, C. (1992) Examining the validity of a scale used in a performance assessment from many angles using the many-facet Rasch model. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA (ERIC Document Reproduction Service, ED 343 442)
- Kesar, O. (1990) Identification and analysis of reading moves in completing a rational deletion cloze. Master's thesis (in Hebrew), School of Education, Hebrew University of Jerusalem, Jerusalem, Israel
- Khalifa, H. (1997) A study in the construct validation of the reading module of an EAP proficiency test battery: validation from a variety of perspectives. Unpublished Ph.D. thesis, University of Reading, England
- Klecka, W.R. (1980) Discriminant Analysis. Quantitative applications in the social science series, No. 19. Thousand Oaks, CA: Sage Publications.
- Krathwohl, D. (1998) Methods of Educational & Social Science Research. New York: Addison-Wesley Educational Publishers
- Kuhn, T. (1970) The Structure of Scientific Revolutions. Chicago: University of Chicago Press
- Kunnan, A. (1998) Approaches to validation in language assessment. In: Kunnan, A. (ed.) Validation in Language Assessment. New Jersey: Lawrence Erlbaum Associates
- Kunnan, A. (1995) Test taker characteristics and test performance: A structural modeling approach. In: Milanovic, M. (ed.) Studies in Language Testing, 2. University of Cambridge Local Examinations Syndicate and Cambridge University Press
- Larson, K. (1981) A study of student test-taking strategies and difficulties. Unpublished manuscript, ESL Section, University of California, Los Angeles
- Lazaraton, A. (2002) A qualitative approach to the validation of oral language tests. In: Milanovic, M. (ed.) Studies in Language Testing, 14. University of Cambridge Local Examinations Syndicate and Cambridge University Press
- Le Blanc, R. and Painchaud, G. (1985) Self-assessment as a second language placement instrument. TESOL Quarterly, 19, pp.673-687
- Li, W. (1992) What is a test testing? An investigation of the agreement between students' test taking processes and test constructors' presumption. Unpublished MA thesis, Lancaster University, England
- Lincoln, Y.S. and Guba, N.K. (1985) Naturalistic Inquiry. Newbury Park, CA: Sage
- Lynch, B. (2001) Rethinking assessment from a critical perspective. Language Testing, 18(4), pp.351-372.
- MacIntyre, P.D. (1995) How does anxiety affect second language learning? A reply to Sparks and Ganschow. The Modern Language Journal, 79, pp.90-99
- MacIntyre, P.D., Noels, K.A. and Clement, R. (1997) Biases in self-ratings of second language proficiency: The role of language anxiety. Language Learning, 47(2), pp.265-287

Madsen, H.S. (1982) Determining the debilitating effects of test anxiety. Language Learning, 32, pp. 133-143

Mathison, S. (1988) Why triangulate? Educational Researcher, 17(2), pp.13-17

McNamara, T. (2001) Language assessment as social practice: challenges for research. Language Testing, 18(4), pp. 333-349

McNamara, T. (1996) Measuring Second Language Performance. New York: Addison Wesley Longman.

McNamara, T. (1990) Item response theory and the validation of an ESP test for health professionals. Language Testing, 7, pp.52-76

Meara, P. (2002) The rediscovery of vocabulary. Second Language Research, 18(4), pp.393-407

Messick, S. (1996) Validity and washback in language testing. Language Testing, 13, pp. 241-256

Messick, S. (1995) Standards of validity and the validity of standards in performance assessment. Educational Measurement: Issues and Practice, 14(4), pp. 5-8

Messick, S. (1989) Validity. In Linn, R. (ed.) Educational Measurement. New York: Macmillan, pp. 13-103

Miles, M and Huberman, A. (1994) Qualitative Data Analysis. London: Sage

Moller, A. (1982) A study in the validation of proficiency tests of English as a foreign language. Ph.D. Thesis. University of Edinburgh.

Moritz, C. (1995) Self-assessment of foreign language proficiency: a critical analysis of issues and a study of cognitive orientations of French learners. Ph.D. Thesis, Cornell University

Moss, P. (1996) Enlarging the dialogue in educational measurement: Voices from interpretive research traditions. Educational Researcher, 25(1), pp.20-28, 43

Moss, P. (1992) Shifting conceptions of validity in educational measurement: Implications for performance assessment. Review of Educational Research, 62(3), pp.229-258

Munby, J. (1978) Communicative Syllabus Design. Cambridge: Cambridge University Press

Myles, L. (1997) The interaction between a computer adaptive test and self-assessment of second language ability. Report on work in progress (Concordia University) at the 19th Annual Language Testing Research Colloquium, Orlando, Florida

Nevo, N. (1989) Test-taking strategies on a multiple-choice test of reading comprehension. Language Testing, 6(2), pp. 199-215

North (1993) The Development of Descriptors on Scales of Proficiency: Perspective, Problems, and a Possible Methodology. Washington, DC: National Foreign Language Center

Nunally, J. (1978) Psychometric Theory, (2nd ed.) New York: McGraw-Hill

- Oatley, K. and Jenkins, J. (1996) Understanding Emotions. Cambridge, MA: Blackwell
- Oller, J.W. (1984) 'g' what is it? In: Hughes, A. and Porter, D. (eds.) Current Developments in Language Testing. London: Academic Press
- Oller, J.W. (1983) Evidence for a general language proficiency factor: an expectancy grammar. In: Oller, J. (ed.) Issues in Language Testing Research. Rowley, M.A.: Newbury House, pp. 3-10
- Oller, J.W. (1979) Language Tests at School: A Pragmatic Approach. London: Longman.
- Oller, J.W. and Perkins, K. (1978) Language proficiency as a source of variance in self-reported affective variables. In: Oller, J.W. and Perkins, K. (eds.) Language in Education: Testing the Tests. Rowley, Mass.: Newbury House
- Omaggio, A. (1986) Teaching Language in Context: Proficiency-Oriented Instruction. Boston, MA: Heinle & Heinle
- O'Malley, M.J. and Chamot, A.U. (1990) Learning Strategies in Second Language Acquisition. Cambridge: Cambridge University Press
- Oscarson, M. (1997) Self-assessment of foreign and second language proficiency. In: Clapham, C. and Corson, D.(eds.) Encyclopedia of Language and Education, Volume 7: Language Testing and Assessment. Netherlands: Kluwer Academic Publishers, pp.175-187
- Oscarson, M. (1989) Self-assessment of language proficiency: Rationale and applications. Language Testing, 6, pp.1-13
- Oscarson, M. (1978) Approaches to Self-Assessment in Foreign Language Learning. Strasbourg, France: Council of Europe, Council for Cultural Cooperation
- Oxford, R. (1990) Language Learning Strategies: What Every Teacher Should Know. New York: Newbury House Publisher
- Pearson, P.D. and Johnson, D.D. (1978) Teaching Reading Comprehension. New York: Holt, Rinehart and Winston
- Pedhazur, E.J. (1973) Multiple Regression in Behavioral Research: Explanation and Prediction. New York: Holt, Rinehart and Winston
- Pekrun, R., Goetz, T., Titz, W. and Perry, R.P. (2002) Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. Educational Psychologist, 37(2), pp.91-105
- Peirce, B. and Stein, P. (1995) Why the "Monkeys Passage" bombed: Tests, genre and teaching? Harvard Educational Review, 65, pp.50-65
- Peirce, B., Swain, M. and Hart, D. (1993) Self-assessment, French immersion and locus of control. Applied Linguistics, 14, pp.25-42
- Phillips, E. M. (1992) The effects of language anxiety on students' oral test performance and attitudes. The Modern Language Journal, 76, pp. 14-26

- Pintrich, P.R. (2000) The role of goal orientation in self-regulated learning. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of Self-Regulation. San Diego, CA: Academic, pp.452-502
- Pressley, M. and Afflerbach, P. (1995) Verbal Protocols of Reading: The Nature of Constructively Responsive Reading. Mahwah, NJ: Lawrence Erlbaum Associates
- Purpura, J. (1999) Learner strategy use and performance on language tests: A structural equation modeling approach. In: Milanovic, M. (ed.) Studies in Language Testing, 8. Cambridge: University of Cambridge Local Examinations Syndicate and Cambridge University Press
- Purpura, J. (1997) An analysis of the relationships between test takers' cognitive and metacognitive use and second language test performance. Language Learning, 47, pp. 289-325
- Purpura, J. (1996) Modeling the relationships between test takers' reported cognitive and metacognitive strategy use and performance on language tests. Ph.D. Thesis, University of California, Los Angeles
- Ready-Morfitt, D. (1991) The role and limitations of self-assessment in testing and research. Paper presented at the 13th Annual Language Testing Research Colloquium, Princeton, New Jersey
- Richardson, L. (2000) Writing: A method of inquiry. In: Denzin, N.K. and Lincoln, Y.S. (eds.) Handbook of Qualitative Research. 2nd ed. London: Sage
- Roizen, M.A. (1982) The population explosion: test-taking strategies project, School of Education, Hebrew University of Jerusalem, course paper
- Ross, S. (1998) Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. Language Testing, 15(1), pp.1-20
- Roth, J.L., Paris, S.G. and Turner, J.C. (2000) Students' perceived utility and reported use of test-taking strategies. Issues in Education, 6(1/2), pp. 67-82
- Sarig, G. (1987) High-level reading in the first and in the foreign language: Some comparative process data. In: Devine, J.; Carrell, P.L. and Eskey, D.E. (eds.) Research in Reading in English as a Second Language. Washington, DC: Teachers of English to Speakers of Other Languages, pp. 105-120
- Savignon, S. (1985) Evaluation of communicative competence: the ACTFL provisional proficiency guidelines. Modern Language Journal, 69, pp.129-134
- Schachter, J. (1990) Communicative competence revisited. In: Harley, B., Allen, P., Cummins, J. and Swain, M. (eds.) The Development of Second Language Proficiency. New York: Cambridge University Press
- Scherer, K.R. (2000) Emotions as episodes of subsystem synchronization driven by other nonlinear appraisal process. In: Lewis, M.D. and Granic, I. (eds.) Emotion, development and self-organization: Dynamic systems approaches to emotional regulation. New York: Cambridge University Press, pp. 70-99
- Schmidt, R. (1990) The role of consciousness in second language learning. Applied Linguistics, 11, pp.17-46.

Schmidt, R., Boraie, D. and Kassabgy, O. (1996) Foreign language motivation: Internal structure and external connections. In: Oxford, R. (ed.) Language Learning Motivation: Pathways to the New Century. Technical Report #11. Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center, pp.9-70

Schmidt, R., Boraie, D. and Kassabgy, O. (1994) Classroom dynamics: Motivation, learning strategies and methods. Paper presented at the TESOL 28th Annual Convention, Baltimore, Maryland

Schunk, D.H. and Ertmer, P.A. (2000) Self-regulation and academic learning: Self-efficacy enhancing interventions. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of Self-Regulation. San Diego, CA: Academic, pp.631-650

Schutz, P.A. and Davis, H.A. (2000) Emotions and self-regulation during test taking. Educational Psychologist, 35(4), pp. 243-256

Schutz, P.A. and DeCuir, J.T. (2002) Inquiry on emotions in education. Educational Psychologist, 37(2), pp. 125-134

Schutz, P.A., Davis, H.A. and Schwanenflugel, P.J. (2002) Organization of concepts relevant to emotions and their regulation during test taking. The Journal of Experimental Education, 70(4), pp. 316-342

Schwandt, T.A. (1994) Constructivist, interpretivist approaches to human inquiry. In: Denzin, N. and Lincoln, Y. (eds.) Handbook of Qualitative Research. London: Sage

Seliger, H.W. (1983) The language learner as linguist: of metaphors and realities. Applied Linguistics, 4, pp. 179-191

Skehan, P. (1991) Progress in language testing: the 1990s. In Alderson, C. and North, B. (eds.) Language Testing in the 1990s: The Communicative Legacy. London: Macmillan, pp. 3-21

Skehan, P. (1984) Issues in the testing of English for specific purposes. Language Testing, 1/2, pp. 202-220

Smith, J.K. (1983) Quantitative versus qualitative research: An attempt to clarify the issue. Educational Researcher, 12, 6-13

Smith, K. and Baldauf, R.B. (1982) The concurrent validity of self-rating with interviewer rating on the Australian Second Language Proficiency Scale. Educational and Psychological Measurement, 42(4), pp.1117-1124

Spolsky, B. (1990) Social aspects of individual assessment. In de Jong, J. and Stevenson, D. (eds.) Individualizing the Assessment of Language Abilities. Clevedon: Multilingual Matters

Spolsky, B. (1967) Do they know enough English? In: Wigglesworth, D. (ed.) Selected Conference Papers of the Association of Teachers of English as a Second Language. Los Altos, CA: Language Research Associates

Tashakkori, A. and Teddlie, C. (1998) Mixed Methodology: Combining Qualitative and Quantitative Approaches. Applied Social Research Methods Series, Volume 46. Thousand Oaks: Sage.

- Thorndike, R.M., Cunningham, G.K., Thorndike, R.L., and Hagen, E.P. (1991) Measurement and Evaluation in Psychology and Education, (5th ed.) New York: Macmillan
- Upshur, J. (1975) Objective evaluation of oral proficiency in the ESOL classroom. In: Palmer, L. and Spolsky, B. (eds.) Papers on Language 1967-1974. Washington, DC: TESOL
- Usher, R. (1996) A critique of the neglected epistemological assumptions of educational research. In: Scott, D. and Usher, R. (eds.) Understanding Educational Research. London: Routledge, pp. 9-32
- Vancouver, J.B. (2000) Self-regulation in organizational settings: A tale of two paradigms. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of self-regulation. San Diego, CA: Academic, pp.303-341
- Vattanapath, R. and Jaiprayoon, K. (1999) An assessment of the effectiveness of teaching test-taking strategies for multiple-choice English reading comprehension tests. Occasional Papers, 8, pp.57-71
- Wall, D. and Alderson, C. (1993) Examining washback: The Sri Lankan impact study. Language Testing, 10, pp.41-69
- Waxman, H.C. and Padron, Y. (1987) The effect of ESL students' perception of their cognitive strategies on reading achievement. Presented at the Annual Meeting of the Southwest Educational Research Association, Dallas.
- Weir, C. (1983) The associated examining board's test in English for academic purposes: An exercise in content validation. In: Hughes, A. and Porter, D. (eds.) Current Developments in Language Testing. London: Academic Press, pp. 147-153
- Weir, C., Huizhong, Y. and Yan, J. (2000) An empirical investigation of the componentiality of L2 reading in English for academic purposes. In: Milanovic, M. (ed.) Studies in Language Testing, 12. Cambridge: University of Cambridge Local Examinations Syndicate and Cambridge University Press
- Wenden, A. (1991) Learner Strategies for Learner Autonomy. Englewood Cliffs, NJ: Prentice-Hall
- Wesche, M. (1993) A comparative study of four placement instruments. Paper presented at the 15th Language Testing Research Colloquium, Cambridge, England and Arnhem, The Netherlands (ERIC Document Reproduction Service No. ED 364 064)
- Wesche, M., Morrison, F., Ready, D. and Pawley, C. (1990) French immersion: Postsecondary consequences for individuals and universities. Canadian Modern Language Review, 46, pp. 43-451
- Wijgh, I. (1996) A communicative test in analysis: Strategies in reading authentic texts. In: Cumming, A. and Berwick, R. (eds.) Validation in Language Testing. Clevedon, UK: Multilingual Matters, pp.154-170
- Wilkinson, L. (1990) SYSTAT: The System for Statistics. Evanston, IL: SYSTAT Inc.
- Wine, J. (1971) Test anxiety and direction of attention. Psychological Bulletin, 76, 92-104
- Yin, R. (1994) Case Study Research. Thousand Oaks: Sage

Zeidner, M. (1998) Test anxiety: The state of the art. New York: Plenum

Zeidner, M. (1995) Adaptive coping with test situations: A review of the literature. *Educational Psychologist*, 30(3), pp. 123-133

Zeidner, M. and Bensoussan, M. (1988) College students' attitudes towards written versus oral test of English as a foreign language. *Language Testing*, 5, pp.100-124

Zeidner, M., Boekaerts, M. and Pintrich, P.R. (2000) Self-regulation: Directions and challenges for future research. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of Self-Regulation. San Diego, CA: Academic, pp.750-768

Zimmerman, B.J. (2000) Attaining self-regulation: A social cognitive perspective. In: Boekaerts, M., Pintrich, P.R. and Zeidner, M. (eds.) Handbook of Self-Regulation. San Diego, CA: Academic, pp.13-39

APPENDICES

APPENDIX 1:
LISTS OF TEST-TAKING STRATEGIES IDENTIFIED
FROM THE LITERATURE

**Cohen's List of Strategies for Taking a Multiple Choice Reading
Comprehension Test (Cohen, 1998:103)**

1. Read the text passage first and make a mental note of where different kinds of information are located.
2. Read the questions a second time for clarification.
3. Return to the text passage to look for the answer.
4. Find the portion of the text that the question refers to and then look clue to the answer.
5. Look for answers to questions in chronological order in the text.
6. Read the questions first so that the reading of the text is directed at finding answers to those questions.
7. Try to produce your own answer to the question before you look at the options that are provided in the test.
8. Use the process of elimination – i.e., select a choice not because you are sure that it is the correct answer, but because the other choices do not seem reasonable, because they seem similar or overlapping, or because their meaning is not clear to you.
9. Choose an option that seems to deviate from the others, is special, is different, or conspicuous.
10. Select a choice that is longer / shorter than the others.
11. Take advantage of clues appearing in other items in order to answer the item under consideration.
12. Take into consideration the position of the option among the choices (first, second, etc.).
13. Select the option because it appears to have a word or phrase from the passage in it – possibly a key word.
14. Select the option because it has a word or phrase that also appears in the question.
15. Postpone dealing with an item or selecting a given option until later.
16. Make an educated guess – e.g., use background knowledge or extratextual knowledge in making the guess.
17. Budget your time wisely on this test.
18. Change your responses as appropriate – e.g., you may discover new clues in another item.

Categorization of Processing Strategies
(Anderson et al, 1999:49)

- I. Supervising strategies is a category which includes strategies in which the reader:
 1. refers to the experimental task;
 2. recognizes loss of concentration;
 3. states failure to understand a portion of the text;
 4. states success in understanding a portion of the text;
 5. adjusts reading rate in order to increase comprehension;
 6. formulates a question;
 7. makes a prediction about the meaning of a word or about text content;
 8. refers to lexical items that impede comprehension;
 9. confirms / disconfirms an inference;
 10. refers to the previous passage; or
 11. responds affectively to text content.

- II. Support strategies is a category which includes strategies in which the reader:
 12. skips unknown words;
 13. expresses a need for a dictionary;
 14. skims reading material for a general understanding;
 15. scans reading material for a specific word or phrase; or
 16. visualizes.

- III. Paraphrase strategies is a category which includes strategies in which the reader:
 17. uses cognates between L1 and L2 to comprehend;
 18. breaks lexical items into parts;
 19. paraphrases;
 20. translates a word or a phrase into the L1;
 21. extrapolates from information presented in the text; or
 22. speculates beyond the information presented in the text.

- IV. Strategies for establishing coherence in text is a category which includes strategies in which the reader:
 23. rereads;
 24. uses context clues to interpret a word or phrase;
 25. reacts to author's style or text's surface structure;
 26. reads ahead;
 27. uses background knowledge;
 28. acknowledges lack of background knowledge; or
 29. relates the stimulus sentence to personal experiences.

- V. Test-taking strategies is a category which includes strategies in which the reader:
30. guesses without any particular considerations;
 31. looks for the answers in chronological order in the passage;
 32. selects an answer not because it was thought to be correct, but because the others did not seem reasonable, seemed similar, or were not understandable;
 33. selects an alternative through deductive reasoning;
 34. matches the stem and /or alternatives to a previous portion of the text;
 35. selects a response because it is stated in the text;
 36. selects a response based on understanding the material read;
 37. makes a reference about time allocation;
 38. reads the questions and options after reading the passage;
 39. reads the questions an options before reading the passage;
 40. changes an answer after having marked one;
 41. receives clues from answering one question that are helpful in answering another;
 42. stops reading the options when they reach the answer;
 43. expresses uncertainty at correctness of an answer chosen;
 44. skips a question and returns to it later;
 45. skips a question that is not understood and leaves the response blank;
 46. marks answers without reading in order to fill the space, or
 47. recognizes during the think-aloud protocol that an answer marked is incorrect.

Test Taking Strategies
(Vattanapath & Jaiprayoon, 1999:70)

Test Taking Strategy	Detail
Strategy 1	Preview or survey the whole test to see how much there is to do.
Strategy 2	Fill your personal details; put your name, candidate number and other details required on your exam sheet.
Strategy 3	Manage your time to ensure that you schedule enough for all portions of the test, use 1-2 minutes to survey the whole test and 10 minutes for checking answers.
Strategy 4	Read the instructions carefully.
Strategy 5	Complete the easy questions first; skip the difficult questions, mark them in some way so that you will remember which ones to come back to later.
Strategy 6	Read all the answers before you choose one.
Strategy 7	Use the process of elimination: eliminate options which are known to be incorrect and choose from among the remaining options.
Strategy 8	Cope with a scoring machine; mark clearly so as not to make more than one mark on any one item.

Test Taking Strategy	Detail
Strategy 9	Never leave an answer blank; answer every question.
Strategy 10	Check your answers to make sure that you have put the answers in the right place and not made careless errors.
Strategy 11	If you do not know the answer, use an intelligent guess.
Strategy 12	Do not change your first answer unless you are convinced that you have made an error.
Strategy 13	Reduce your test anxiety.
Strategy 14	Quickly skim the questions first, then read the passage carefully and critically.
Strategy 15	Read and analyze questions carefully; underline the key words such as not, except, incorrect and false. These key words give you clues to the correct answer.
Strategy 16	Be sure to answer the questions only on the basis of the information given to you in the passage, and not from outside information you may happen to know.
Strategy 17	Notice whether a detail question refers to a specific line, sentence, or quotation from the reading passage. Use key words in the reading passage. The answer to such a question is almost certain to be found on or near the reference in the passage.
Strategy 18	Use the context clues.
Strategy 19	If there are several choices that seem correct on a multiple-choice question, select the most complete answer, that is, the one that includes the other choices.
Strategy 20	Learn how to handle special choices. Items that include "all of the above" or "none of the above" choices can be answered easily if you convert each choice into a true-false statement. When you are certain that at least two choices are correct on a multiple-choice item, select the answer "all of the above". If there are at least two incorrect choices, select "none of the above". Be sure to read each choice carefully.
Strategy 21	Try to anticipate the answer even though it is difficult. Then look for it among the choices. If it is not there, forget it. You must choose the best answer available.
Strategy 22	If you do not have enough time to read the whole passage, read the topic sentence which is found most often in the first sentence of every paragraph.

Positive and Negative Strategies

(Roth et al, 2000:77)

Positive Strategies

1. Read story first, then read questions and look for answers in story.
2. Look at other questions for clues to questions you don't understand.
3. Watch the time.
4. Answer questions you know first.
5. Move away from student who distracts you.
6. Use process of elimination.
7. Rely on prior knowledge of the subject.
8. Prepare for test by comparing strategies with other students.
9. Go back to check answers if you finish early.
10. Enjoy tests because they are a challenge.
11. I answer the easy questions first and then go back to the hard ones.
12. I make a mark on the item that I want to go back to.
13. I read the directions carefully.
14. I go back and check my answers when I finish the test.
15. I try to narrow down the choices to find the best one.
16. I try to think about possible answers to the question before I read the choices.
17. I read all the possible answer choices before choosing one.
18. I read the entire passage carefully before I answer the questions.
19. I check my answer sheet to be sure I put my answers in the right places.
20. When I finish the test, I go back and read the passages again to check my answers.
21. I think about what the questions are asking before I read the passage.
22. When the answer is not in the passage, I use clues from the passage and my own experiences to answer the question.
23. I think about the topic before I read the passage.
24. I try to concentrate when I take the test.
25. I watch the time carefully so I know how much time I have left.

Negative Strategies

1. Get tired and start filling in bubbles without reading the questions.
2. Read questions first, then look for answers without really reading passage.
3. Get nervous and get stuck on one question for a long time.
4. Forget about time limits.
5. Read and answer questions quickly to be first in the class to finish.
6. Have trouble concentrating, so you look around the room.
7. Just guess on questions that are confusing.
8. If finished early, don't check your answers.
9. Dread taking tests because you know more than the test will show.
10. Look at someone else's answer sheet to check answers to difficult questions.
11. I try to ask the teacher for hints or clues about the right answer.
12. I get tired of the test and just fill in answers without thinking.
13. I do not look back at the passage when I am answering the questions.
14. I run out of time and don't get to try all the questions.
15. I fill in the answers without reading the passage.

16. I look for the answers to the questions without reading the passage.
17. I try to finish the test as fast as I can.
18. I look at other students' answers during the test.
19. I skip hard words in the passage without trying to figure out what they mean.
20. I have trouble concentrating on the test.
21. I get so nervous during a test I forget the things I usually know.
22. When I come to a hard question, I don't go on until I finish.
23. I lose my place on the answer sheet.
24. I watch to see when other students finish the test.
25. It is hard for me to sit still for the whole test.

APPENDIX 2:
SEMI-STRUCTURED INTERVIEW: STUDENT INTERVIEW PROTOCOL

1. Why do you want to study English?
2. Where do you use / are you going to use English?
3. Describe the difficulties you face in using English.
4. Describe the language of someone who is at the top level of proficiency.

APPENDIX 3:
OPEN-ENDED QUESTIONNAIRE: TEACHERS' QUESTIONNAIRE

I am currently investigating Egyptian adult learners' perceptions about their own competence in EFL and their expectations in terms of terminal language objectives. I would very much appreciate your input regarding these issues by answering the following questions. Thank you very much for your time and effort.

1. Why do you think Egyptian adults study English? What are the different reasons that bring them to CACE?

2. What are the different contexts or situations where your students would use English?

3. Please describe the language of a person at the “top” of the English proficiency continuum?

APPENDIX 4: **TEST TAKING STRATEGIES QUESTIONNAIRE (TTSQ)**

The researcher is conducting this survey in order to find out the different strategies you used while taking the test and your reactions to this test. The information we obtain from these questionnaires will be used in the planning and design of future placement tests in the Center for Adult & Continuing Education. Your thoughtful completion of this questionnaire would be most appreciated. Thank you for your time and cooperation.

Name: _____

SID: _____

Rate each of the following statements by **circling** a number on the following scale that best reflects your opinion in accordance with the descriptors:

Rating	Description
3	I used this strategy SEVERAL times during the test.
2	I used this strategy only ONCE or TWICE during the test.
1	I did NOT use this strategy during the test.

#	STATEMENT	RATING		
		3	2	1
1	I read the questions and options before choosing one.			
2	I stopped reading options when I got to the one that seemed correct.			
3	I translated the question and the options.			
4	I selected the option by eliminating the other 3 options.			
5	I inserted each option one at a time in the question.			
6	I made an educated guess using background knowledge.			
7	I tried to produce my own answer to the question before looking at the options provided.			
8	I reread the questions and options for clarification.			
9	I postponed dealing with a question or selecting a given option until later.			
10	If there was a question that I did not understand, I left the answer blank.			
11	I guessed without any particular considerations.			
12	I changed my answer when appropriate.			
13	When I was not sure of the answer, I selected an option that was longer / shorter than the others.			
14	When I was not sure of the answer, I looked for an option that seemed to be different from the others.			
15	I ran out of time without trying all the questions.			

Rating	Description
3	I used this strategy SEVERAL times during the test.
2	I used this strategy only ONCE or TWICE during the test.
1	I did NOT use this strategy during the test.

#	STATEMENT	RATING		
		3	2	1
16	I tried to finish the test as fast as possible.			
17	I started by reviewing or surveying the whole test.			
18	I monitored the time.			
19	I watched to see when other students finish the test.			
20	I read the instructions carefully.			
21	I left an answer blank.			
22	I went back and reviewed or checked my answers.			
23	I got stuck on one question for a long time.			
	Strategies I Used in the Writing Section			
24	I read the prompt in English only.			
25	I read the prompt first in Arabic and then in English.			
26	I read the prompt first in English and then in Arabic.			
27	I read the prompt in Arabic only.			
28	I reread the prompt.			
29	I formulated my ideas first in Arabic and then translated them into English and wrote them down.			
30	I used words from the prompt in my answer.			
31	I formulated my ideas in English only.			
32	I rewrote words or phrases in my answer.			
	Strategies I Used in the Listening Section			
33	I had sufficient time to read the options of the next question before listening to the question on tape.			
34	I used my knowledge of grammar to answer the question.			

Rating	Description
3	I used this strategy SEVERAL times during the test.
2	I used this strategy only ONCE or TWICE during the test.
1	I did NOT use this strategy during the test.

#	STATEMENT	RATING		
		3	2	1
	Strategies I Used in the Reading Section			
35	I read the passage first.			
36	I read the questions first before reading the passage.			
37	I reread the whole passage.			
38	I reread parts of the passage.			
39	I translated relevant parts of the passage to understand.			
40	I read the passage and then summarized the ideas in Arabic.			
41	I was able to guess the meaning of an unknown word from the context.			
42	I skipped unknown words.			
43	I pronounced or sounded out the word to find its meaning.			
44	I used my knowledge of grammar to answer the question.			
45	I looked for the portion of the passage that the question refers to and then looked there for clues to the answer.			
46	I matching material from the passage with material in the question and in the options to find the answer.			
47	I selected an option because it appeared to have a word or phrase from the passage in it.			
48	I selected an option based on understanding the passage I read.			
49	I got clues from answering one question that were helpful in answering another question.			

Did you use a strategy while taking the test that was not mentioned in this questionnaire?

☐ Yes ☐ No

If you answered YES, what was the strategy and did you use it once or twice or more?

Rate each of the following statements by circling a number on the following scale that best reflects your opinion in accordance with the descriptors:

Rating	Description
6	I STRONGLY AGREE.
5	I AGREE.
4	I SLIGHTLY AGREE.
3	I SLIGHTLY DISAGREE.
2	I DISAGREE.
1	I STRONGLY DISAGREE.

#	STATEMENT	RATING					
		6	5	4	3	2	1
1	I tried my hardest on this test.						
2	I thought I did well.						
3	I thought that the test was easy.						
4	I felt so anxious that I wanted to get the answers from another person.						
5	I felt that the test was confusing.						
6	I thought that the test was difficult.						
7	I felt prepared to take this test.						
8	I knew what to do during the test.						
9	I thought I knew why it was important to do my best.						
10	I had enough time to finish the test.						
11	I felt that taking the test was a big challenge.						
12	I felt a sense of achievement after completing the test.						
13	I felt relieved when the test was over.						
14	I felt nervous or anxious during the test.						
15	I had difficulty in concentrating during the test.						
16	I kept looking around the testing room during the test.						
17	I dread taking tests in general.						
18	I dread taking tests because I know that they do not show my true ability.						
19	I felt tired during the test.						
20	I gave up because the test was too difficult.						

Rate each of the following statements by circling a number on the following scale that best reflects your opinion in accordance with the descriptors:

Rating	Description
6	I STRONGLY AGREE.
5	I AGREE.
4	I SLIGHTLY AGREE.
3	I SLIGHTLY DISAGREE.
2	I DISAGREE.
1	I STRONGLY DISAGREE.

#	STATEMENT	RATING					
		6	5	4	3	2	1
21	I got tired and so I started answering without reading the question.						
22	I forgot the things I usually know because I was so nervous.						
23	I was sure of the correct answer in most questions.						
24	I felt that the test was interesting.						
25	I felt bored while taking the test.						
26	Taking the test was a pleasant experience.						
27	I was frustrated because I did not have enough time.						
28	I think it is a good idea to have the instructions in both Arabic and English.						
29	I think it is a good idea to have the writing prompt in both Arabic and English.						

Briefly describe your feelings during the test.

If you felt anxious during the test, in what section did you experience this feeling?

Did this feeling of anxiety differ from one section of the test to another?



استقصاء حول أساليب الإجابة على أسئلة الاختبار

عزيزي الدارس

يهدف الباحث من إجراء هذه الدراسة إلى التعرف على الأساليب التي اتبعتها أثناء إجابتك على أسئلة الاختبار وكذا ردود أفعالك تجاهه وسوف نستخدم المعلومات التي نحصل عليها من هذه الاستقصاءات في تخطيط وتصميم اختبارات تحديد المستوى اللغوي مستقبلاً بمركز تعليم الكبار والتعليم المستمر.

ونحن نتوجه إليك بالشكر على وقتك وتعاونك ونقدر إكمالك لهذا الاستقصاء حتى نهايته.

الاسم :

رقم الدارس :

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها :

الرقم	المعنى/الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب.

التقدير			العبارة
١	٢	٣	
١	٢	٣	١- قمت بقراءة الأسئلة والاختيارات قبل أن أختار إحداها.
١	٢	٣	٢- توقفت عن قراءة الاختيارات عندما توصلت للاختيار الذي بدأ لي صحيحاً.
١	٢	٣	٣- قمت بترجمة السؤال واختياراته.
١	٢	٣	٤- اخترت الإجابة الباقية بعد أن استبعدت ثلاثة اختيارات.
١	٢	٣	٥- جرب استخدام كل اختيار مع السؤال بشكل منفصل.
١	٢	٣	٦- لجأت للتخمين الذكي على أساس المعلومات المتاحة لي.
١	٢	٣	٧- حاولت الإجابة على السؤال قبل النظر إلى الاختيارات الواردة.

الرقم	المعنى/ الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب.

التقدير			العبارة
١	٢	٣	
١	٢	٣	٨- أعدت قراءة السؤال والاختيارات للاستيضاح.
١	٢	٣	٩- كنت أؤجل الإجابة على السؤال أو اختيار الإجابة لوقت لاحق.
١	٢	٣	١٠- إذا كان هناك سؤالاً لم أفهمه ، فلم أعط له أية إجابة.
١	٢	٣	١١- قمت بعملية تخمين عشوائي.
١	٢	٣	١٢- غيرت إجابتي عند اللزوم.
١	٢	٣	١٣- عندما لم أكن متأكدًا من الإجابة اخترت الإجابة الأطول أو الأقصر من بقية الإجابات.
١	٢	٣	١٤- عندما لم أكن متأكدًا من الإجابة بحثت عن الاختيار الذي كان يبدو مختلفًا عن الاختيارات الأخرى.
١	٢	٣	١٥- نفذ الوقت مني قبل انتهائي من قراءة جميع الأسئلة.
١	٢	٣	١٦- حاولت الانتهاء من إجابة الاختبار بأسرع ما يمكن.
١	٢	٣	١٧- بدأت بقراءة سريعة للاختبار بأكمله.
١	٢	٣	١٨- كنت أراجع الوقت الذي انقضى والوقت المتبقي أثناء الإجابة.
١	٢	٣	١٩- كنت أرقب الآخرين لأرى متى ينتهون من الاختبار.
١	٢	٣	٢٠- قرأت التعليمات جيداً.
١	٢	٣	٢١- تركت سؤالاً بغير إجابة.
١	٢	٣	٢٢- قمت بمراجعة إجاباتي.
١	٢	٣	٢٣- توقفت طويلاً عند أحد الأسئلة.
			اساليب استخدمتها للإجابة على جزء الكتابة :
١	٢	٣	٢٤- قرأت رأس السؤال باللغة الإنجليزية فقط.
١	٢	٣	٢٥- قرأت رأس السؤال باللغة العربية - أولاً ثم بالإنجليزية.
١	٢	٣	٢٦- قرأت رأس السؤال باللغة الإنجليزية - أولاً ثم بالعربية.
١	٢	٣	٢٧- قرأت رأس السؤال باللغة العربية فقط.
١	٢	٣	٢٨- أعدت قراءة رأس السؤال مرة أخرى.
١	٢	٣	٢٩- بدأت بصياغة أفكارى باللغة العربية ثم ترجمتها إلى الإنجليزية وقمت بكتابتها.
١	٢	٣	٣٠- استخدمت في الإجابة على سؤال الكتابة كلمات من التي وردت في رأس السؤال.
١	٢	٣	٣١- قمت بصياغة أفكارى باللغة الإنجليزية فقط.
١	٢	٣	٣٢- أعدت كتابة بعض الكلمات أو العبارات في إجابتي على سؤال الكتابة.

الرقم	المعنى/الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب.

التقدير			العبارة
١	٢	٣	
			أساليب استخدمتها للإجابة على أسئلة الاستماع :
١	٢	٣	٣٢- كان لدي الوقت الكافي لقراءة اختيارات السؤال التالي قبل الاستماع إليه من الشريط.
١	٢	٣	٣٤- استخدمت معلوماتي عن القواعد في إجابة السؤال.
			أساليب استخدمتها للإجابة على أسئلة القراءة :
١	٢	٣	٣٥- بدأت بقراءة القطعة.
١	٢	٣	٣٦- قمت بقراءة الأسئلة قبل قراءة القطعة.
١	٢	٣	٣٧- أعدت قراءة القطعة بأكملها مرة أخرى.
١	٢	٣	٣٨- أعدت قراءة أجزاء من القطعة مرة أخرى.
١	٢	٣	٣٩- ترجمت أجزاء معينة حتى أتمكن من فهم القطعة.
١	٢	٣	٤٠- قرأت القطعة ثم لخصت الأفكار باللغة العربية.
١	٢	٣	٤١- استطعت تخمين معنى كلمة لا أعرفها من السياق.
١	٢	٣	٤٢- لم أتوقف عند الكلمات التي لم أعرفها.
١	٢	٣	٤٣- كنت أنطق بالكلمة للوقوف على معناها.
١	٢	٣	٤٤- استخدمت معلوماتي عن القواعد في إجابة السؤال.
١	٢	٣	٤٥- بحثت عن الجزء الذي له علاقة بالسؤال للحصول على مفاتيح الإجابة.
١	٢	٣	٤٦- ربطت بين ماورد في القطعة وما ورد في السؤال وما ورد في الإختبارات للإجابة على السؤال.
١	٢	٣	٤٧- لجأت إلى أحد الإختيارات كإجابة لأنه كان يتضمن كلمة أو عبارة من القطعة.
١	٢	٣	٤٨- قام اختياري على أساس فهمي للقطعة التي قرأتها.
١	٢	٣	٤٩- ساعدتني إجابة أحد الأسئلة في إجابة سؤال آخر.

* هل هناك أي أسلوب استخدمته أثناء إجابتك على هذا الإختبار لم يذكر في هذا الإستقصاء ؟

نعم ☐ لا ☐

إذا كانت الإجابة نعم من فضلك بين هذا الأسلوب وهل استخدمته مرة أو مرتين أم أكثر ؟

.....

.....

.....

.....

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها :

الرقم	المعنى/الشرح
٦	أوافق تماماً
٥	أوافق
٤	أوافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

التقدير						العبارة
١	٢	٣	٤	٥	٦	
١	٢	٣	٤	٥	٦	١- بذلت قصاري جهدي في هذا الاختبار.
١	٢	٣	٤	٥	٦	٢- أعتقد أنني أجبت إجابة جيدة.
١	٢	٣	٤	٥	٦	٣- اعتقد أنه كان اختباراً سهلاً.
١	٢	٣	٤	٥	٦	٤- شعرت بالضيق لدرجة الرغبة في حصولي على الإجابة من شخص آخر.
١	٢	٣	٤	٥	٦	٥- شعرت بأن الإختبار قد أصابني بالإرتباك.
١	٢	٣	٤	٥	٦	٦- اعتقد أنه كان اختباراً صعباً.
١	٢	٣	٤	٥	٦	٧- كنت مستعداً لهذا الاختبار.
١	٢	٣	٤	٥	٦	٨- كنت أعرف ما ينبغي عليّ عمله أثناء الإختبار.
١	٢	٣	٤	٥	٦	٩- أعتقد أنني كنت أعرف أهمية أن أبذل قصاري جهدي.
١	٢	٣	٤	٥	٦	١٠- كان لدي الوقت الكافي للانتهاء من الاختبار.
١	٢	٣	٤	٥	٦	١١- كان الإختبار بالنسبة لي تحدياً كبيراً.
١	٢	٣	٤	٥	٦	١٢- شعرت بالإنجاز عندما أكملت الإجابة على الاختبار.
١	٢	٣	٤	٥	٦	١٣- أحسست بالإرتياح عندما انتهيت من الاختبار.
١	٢	٣	٤	٥	٦	١٤- كنت متوتراً أثناء الاختبار.
١	٢	٣	٤	٥	٦	١٥- وجدت صعوبة في التركيز أثناء الاختبار.
١	٢	٣	٤	٥	٦	١٦- تجولت ببصري كثيراً في قاعة الاختبار أثناء الإجابة.
١	٢	٣	٤	٥	٦	١٧- أخشى الإختبارات عامة.
١	٢	٣	٤	٥	٦	١٨- أخشى الإختبارات لأنني أعرف أنها لا تظهر كفايتي كاملة.
١	٢	٣	٤	٥	٦	١٩- شعرت بالإجهاد أثناء الاختبار.
١	٢	٣	٤	٥	٦	٢٠- لم أكمل الاختبار لأنه كان شديد الصعوبة.

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها :

الرقم	المعنى/ الشرح
٦	لوافق تماما
٥	لوافق
٤	لوافق إلى حد ما
٣	شبه موافق إلى حد ما
٢	شبه موافق
١	شبه موافق على الإطلاق

النتيجه						العبارة
١	٢	٣	٤	٥	٦	
١	٢	٣	٤	٥	٦	٢١- شعرت بالإجهاه وبدأت الإجابة عشوائيا بدون قراءة الأسئلة.
١	٢	٣	٤	٥	٦	٢٢- نسيت بعض الأشياء التي أعرفها بسبب توترى.
١	٢	٣	٤	٥	٦	٢٣- كنت متأكدا من الإجابة الصحيحة في أغلب الأسئلة.
١	٢	٣	٤	٥	٦	٢٤- شعرت أنه كان اختبار مشوقاً.
١	٢	٣	٤	٥	٦	٢٥- شعرت بالملل أثناء الاختبار.
١	٢	٣	٤	٥	٦	٢٦- كان الاختبار تجربة سارة.
١	٢	٣	٤	٥	٦	٢٧- أصابني الإحباط بسبب ضيق الوقت.
١	٢	٣	٤	٥	٦	٢٨- أعتقد أنها فكرة جيدة أن تكون التعليمات باللغتين العربية والإنجليزية.
١	٢	٣	٤	٥	٦	٢٩- أعتقد أنها فكرة جيدة أن يكون رأس السؤال في جزء الكتابة باللغتين العربية والإنجليزية.

* في كلمات مختصرة صف شعورك أثناء أدائك للاختبار ؟

* في حالة شعورك بالتوتر أثناء الإختبار في أي جزء كان هذا الشعور؟

• هل اختلف هذا الشعور من جزء لآخر؟

APPENDIX 5:
TRANSCRIPTS OF THE PILOT STUDENT INTERVIEW PROTOCOL
Tapescript of Pilot Interview #1 - Male, 35+, Basic Level
(Translated)

R = Researcher and I = Interviewee

Researcher explained purpose of interview and obtained permission to tape record the session.

R: Why are you studying English?

I: I am an agricultural engineer and I work in a foreign company. I was placed in an embarrassing situation. A foreigner came to the office and handed me work to do. I had to understand all the information very well which was difficult for me. I was put in a difficult situation where I had to look for someone good at English to help me. This led me to think about studying English, so I came to apply. I sat for the placement test and I was placed in level 2. I got 66% percent on the end-of-term test. I did not consider this a good grade because I was not studying the language casually but I intended to use it and I want to improve it. I attended level 3 but I did not have enough time to study. At the end of the term I also got 68%. So with these two grades, 66 and 68, I felt that I was not doing well or as well as I want.

R: What are the situations where you use English?

I: At work when dealing with the foreign experts.

R: What are the most important skills you need in English?

I: All. I have to listen well. I have to write well also. The main obstacle is to write. Even during the placement test, when I first applied, when asked to write 5 sentences, I left this question.

R: You didn't try?

I: I left this question. I couldn't do it. In level 3 I felt there was some improvement, but not as much as I wanted.

R: What is the level of proficiency you are aiming to achieve?

I: To speak well.

R: Like whom? Can you describe this person?

I: Give you names?

R: No, I mean, describe the language you are aiming for at the end.

I: To know the language very well, read well, write well. Reading is important for an engineer because I need to read certain terminology. It was very embarrassing for me to have to bring someone to explain to me what was said, especially when I am this person's supervisor.

R: How long do you think you need to reach your goal?

I: How much time? The problem is that I'm studying while I'm working.

R: Let's assume that you will continue in the same pattern. How many years will it take you or do you need?

I: 2-3 years because I'm very keen to learn.

R: Please evaluate yourself in the four skills: listening, speaking, reading and writing.

I: How?

R: On a scale of 1-10.

I: Reading - 6 or 7 out of 10. I read well. I write well but it is less than reading. Speaking is 5 out of 10 and listening is less than speaking. It is the most difficult because I need practice in the classroom, through cassettes. Grammar is good.

R: Thank you very much for your time.

Tapescript of Pilot Interview #2 - Male, 23-35, Advanced Level

R = Researcher and I = Interviewee

Researcher explained purpose of interview and obtained permission to tape record the session.

R: Could I start from the beginning? I know a bit about why you're studying English, but could you repeat it again for me?

I: Actually, it is because of two reasons that there is a Human Resources & Training department in my company that qualifies everybody to raise their standard and also I'm looking for my personal situation, or let's say, I'm looking for having after some time the TOEFL exam because ...

R: You are?

I: Yes, because I will apply to have my Masters and my PhD in the States so I should have the TOEFL exam for a higher degree. I dream to have a score of more than 600 or from 550 to 600 so this should take from me a lot of studying. So when they announced about this in my company, I applied for it and that's it. .. but maybe three years ago I was in college and I applied also for studying here in the AUC two times, but there was one year separating the two times. The first time I was in level 11 and the second time I was in level 18 and those times my teacher was Dr. Nemat Matta - she is a very, very good teacher.

R: Yes!

I: and it was my luck that I had Mrs. Nemat in those times.

R: Yes, yes.

I: and right now, I'm starting again.

R: Conversation?

I: Conversation. Yes. I felt that I need to study conversation also because I have been in the States last winter.

R: For how long were you in the States?

I: For 5 weeks.

R: Training?

I: Training, yes but not language training. I was training for some new machines.

R: Yes, yes.

I: For me it was easy communicating with the people there but I felt that I am missing something also that's why I am starting here.

R: You felt you missed something. What does that mean?

I: This means when I was watching TVs that I hear some words that I couldn't understand.

R: Here in Cairo or in the States.

I: In the States.

R: Yes.

I: Also a few words that when I was talking with people when they say something I could them what is the meaning of this or something because I couldn't understand from the first time I have heard that's why and also when there's something that they're talking in their social life or politics or something new. There's some expression they use or something like that that I don't understand it - that's why.

R: Which is American?

I: Yes.

R: Do you use English now in your work?

I: A little bit. Yes when I deal with some foreign people because we used to have a lot of foreign people in the company installing new machines for doing some researches for doing some joint venture work. I deal with a lot of Germans, Koreans, some Americans, we are dealing with USAID. They have some programs which have people when they retire they can work for free. We can afford their expenses living here and so we have 2 or 3 experts in some different subjects one of them was in the quality assurance, one of them was in the human resources, and I think the third one was for computers. And although they are old but they have lived each one of them was staying for 2 or 3 months.

R: Have you ever been in a situation that you found difficult, an English language situation that was difficult for you?

I: We're talking about real difficulty, no, but sometimes I couldn't find the exact work that I'd like to say - this happens.

R: Yes, this happens and in reading or writing, did you ever face that?

I: A little bit. A little bit. I can't say it's rare, it happened, yes.

R: I'd like to go back, you said you want to reach the level over 600 in TOEFL - can you describe someone who is fluent or who is at the highest stage of proficiency? I mean you feel you are not right there yet, can you describe "there", that level you want to reach, the language? What does "there" mean?

I: I describe who was born there, born in a country whose mother tongue is English. I would say that he was educated in English in everything, he was not speaking in Arabic at all. And I had some friends, they were born abroad maybe and they came back here and they were with me in the college and you know I can tell it from the accent - sometimes when I hear the accent, I can say that these group speak English very well, speak English perfect.

R: And is that what you want to reach?

I: What I want to reach is that when I speak with someone in any subject I could speak with him, I could understand everything he says, that's when I speak or listen to the radio, the VOA or when I see something on the CNN or something, I like to understand every word said.

R: Well thank you very much for your time and the best of luck.

**APPENDIX 6: PRELIMINARY VERSION OF THE
TEST-TAKING STRATEGIES QUESTIONNAIRE**

The researcher is conducting this survey in order to find out the different strategies you used while taking the test and your reactions to this test. The information we obtain from these questionnaires will be used in the planning and design of future placement tests in the Center for Adult & Continuing Education. Your thoughtful completion of this questionnaire would be most appreciated. Thank you for your time and cooperation.

Name: _____

Rate each of the following statements by **circling** a number on the scale that best reflects your opinion.

3	=	I used this strategy SEVERAL times during the test.
2	=	I used this strategy only ONCE or TWICE during the test.
1	=	I did NOT use this strategy.

1.	I read the questions and options before choosing one.	3	2	1
2.	I stopped reading options when I got to the one that seemed correct.	3	2	1
3.	I translated the question and the options.	3	2	1
4.	I selected an option by eliminating the other 3 options.	3	2	1
5.	I inserted each option one at a time in the question.	3	2	1
6.	I made an educated guess using background knowledge.	3	2	1
7.	I tried to produce my own answer to the question before looking at the options provided.	3	2	1
8.	I reread the questions and options for clarification.	3	2	1
9.	I postponed dealing with a question or selecting an option until later.	3	2	1
10.	I skipped a question that I did not understand and left the answer blank.	3	2	1
11.	I guessed without any particular considerations.			
12.	I changed my answer when appropriate.	3	2	1
13.	I selected an option that is longer / shorter than the others.	3	2	1
14.	I looked for an option that seemed to be different from the others.	3	2	1
15.	I ran out of time without trying all the questions.	3	2	1
16.	I tried to finish the test as fast as possible.	3	2	1
17.	I started by previewing or surveying the whole test.	3	2	1
18.	I monitored the time.	3	2	1
19.	I watched to see when other students finish the test.	3	2	1
20.	I read the instructions carefully.	3	2	1
21.	I never leave an answer blank.	3	2	1
22.	I went back and reviewed or checked my answers.	3	2	1
23.	I got stuck on one question for a long time.	3	2	1

3	=	I used this strategy SEVERAL times during the test.
2	=	I used this strategy only ONCE or TWICE during the test.
1	=	I did NOT use this strategy.

Specific Strategies Related to the Writing Section

24.	I read the prompt in English only.	3	2	1
25.	I read the prompt first in Arabic and then in English.	3	2	1
26.	I read the prompt first in English and then in Arabic.	3	2	1
27.	I read the prompt in Arabic only.	3	2	1
28.	I reread the prompt.	3	2	1
29.	I formulated my ideas first in Arabic and then translated them into English and wrote them down.	3	2	1
30.	I used words from the prompt in my writing.	3	2	1
31.	I formulated my ideas in English only.	3	2	1
32.	I rewrote words or phrases.	3	2	1

Specific Strategies for Listening

33.	I read the options of the next question before listening to the question on tape.	3	2	1
34.	I used knowledge of grammar to answer the question.	3	2	1

Specific Strategies for Reading

35.	I read the passage first.	3	2	1
36.	I read the questions first before reading the passage.	3	2	1
37.	I reread the whole passage.	3	2	1
38.	I reread parts of the passage.	3	2	1
39.	I translated relevant parts of the passage to understand.	3	2	1
40.	I read the passage & then summarized the ideas in Arabic.	3	2	1
41.	I guessed the meaning of a word from the context.	3	2	1
42.	I skipped unknown words.	3	2	1
43.	I pronounced or sounded out the word to find its meaning.	3	2	1
44.	I used knowledge of grammar to answer the question.	3	2	1
45.	I looked for the portion of the passage that the question refers to and then looked there for clues to the answer.	3	2	1
46.	I matched material from the passage with material in the question and in the options.	3	2	1
47.	I selected an option because it appeared to have a word or phrase from the passage in it.	3	2	1
48.	I selected an option based on understanding the passage I read.	3	2	1
49.	I got clues from answering one question that were helpful in answering another question.	3	2	1

Rate each of the following statements by marking the box on the scale that best reflects your opinion.

- 6

=

I STRONGLY AGREE.
- 5

=

I AGREE
- 4

=

I SLIGHTLY AGREE
- 3

=

I SLIGHTLY DISAGREE
- 2

=

I DISAGREE
- 1

=

I STRONGLY DISAGREE

#	STATEMENT	RATING SCALE					
		6	5	4	3	2	1
1	I tried my hardest on this test.						
2	I thought I did well.						
3	I thought the test was easy.						
4	I thought the test was difficult.						
5	I thought the test was confusing.						
6	I felt like cheating.						
7	I felt prepared to take this test.						
8	I knew what to do during the test.						
9	I felt I knew why it was important to do my best.						
10	I had enough time to finish the test.						
11	I enjoyed taking the test.						
12	I thought taking the test was a challenge.						
13	I felt relieved when the test was over.						
14	I felt nervous during the test.						
15	I had difficulty in concentrating.						
16	I kept looking around the testing room during the test.						
17	I dread taking tests because I know more than the test will show.						
18	I felt tired during the test.						
19	I gave up because the test was too difficult.						
20	I got tired and started answering without reading the question.						
21	Because I was so nervous I forgot the things I usually know.						
22	I was not sure of the correct answer in most questions.						
23	I thought the test was interesting.						
24	I felt bored while taking the test.						
25	Taking the test was a pleasant experience.						
26	I was frustrated because I did not have enough time.						
27	I think it is a good idea to have the instructions in both Arabic and English.						
28	I think it is a good idea to have the writing prompt in both Arabic and English.						

APPENDIX 7:

RECORD OF THE 36 STUDENT INTERVIEW PROTOCOLS

RECORD OF INTERVIEW 1: FEMALE, 35+, ADVANCED

Reason(s) for studying English

To finish her Master's degree in mass communication.

Use of English

As a wife of a diplomat she uses English very often to communicate with other foreigners and friends. She works in the Ministry of Culture and rarely uses English there. She uses English in her social life and sometimes with her children because she sometimes expresses more in English using some words and sentences. She sometime uses English so as not to let other people understand what she's saying to her children.

Difficulties in four skills

She studied at school and some courses during university but was not sufficient. She prefers reading articles and does not find reading novels difficult but she has no time. She has forgotten her writing skills since she had left university 22 years ago. She needs practice in spelling especially because of the silent letters in the words. She finds it difficult to find the right word. It is not difficult to write sentences but has difficulty in words. She finds speaking easy in social occasions and did not remember a difficult situation. She could not recall a difficult situation for listening. She stated that as long as she is concentrating she could understand very well, especially TV. On rare occasions she would not understand some word. She hopes to be able to remember how to write a complete sentence with the correct grammar and spelling. She does not need to learn how to speak and is looking specifically for improving her writing.

Description of "proficiency in English"

Speaks fluent English. Speaks without thinking and can express herself very quickly. Doesn't stop to think or to translate from her own language. Expresses herself quickly without even thinking. Speaks and uses simple and correct language - not complicated words or expressions.

RECORD OF INTERVIEW 2: MALE, 35+, LOWER INTERMEDIATE

Reason(s) for studying English

As a cashier, English is important to him because most accounting jobs nowadays depend on the language, especially currency exchange, external relations and for computer studies. Computers are important in a bank and therefore, the language is very important for him in this field. Language also helps in solving the problems of colleagues and stated that in the Dokki branch where he works, there are a lot of foreign clients and his job is not only on the computer but also to solve problems with clients.

Use of English

He uses English at work, especially with computers. Dealing with clients occurs occasionally, not as a basic part of his job. He reads English on the screens and in machine manuals. The Help menu is in English. He is also doing higher studies, a diploma in information systems and needs English.

Difficulties in four skills

The difficulty he faces is in correct pronunciation. Reading is no problem. The problem in listening is the speed. He could not follow a conversation. He has no problem in writing and spelling and everything is very good. He hopes to improve speaking and listening, focusing on speaking.

Description of "proficiency in English"

Can understand everything very well and can even understand the inner meaning of words with different meanings. In Arabic one can understand multiple meanings of an expression and to be proficient is to reach this stage in English. It is not just the meaning of the word, but to understand the sentence with different meanings. The person at the top, doesn't take long in understanding and speaks well.

RECORD OF INTERVIEW 3: MALE, 23, BASIC

Reason(s) for studying English

He is a graduate of the Institute of Koranic Reading. A graduate travels to different foreign countries to teach and therefore, needs the language to explain the meaning of the Koran. He stressed that the Koran is in Arabic but the meaning is in English. He stated that he also wanted to teach computer.

Use of English

He does not really use the language and whenever he finds something in English he tries to read it.

Difficulties in four skills

He finds reading and translation difficult. His writing is good since he can write words. He stated that his listening and speaking are weak. He hopes to achieve complete mastery of English and to reach the highest level in English in all the 4 skills.

Description of "proficiency in English"

Pronunciation of words and translation ability.

RECORD OF INTERVIEW 4: MALE, 35+, INTERMEDIATE

Reason(s) for studying English

Because it is very modern and he was late. He studied engineering and should have studied English along with it.

Use of English

He does not use English because he works in a company, implementing the underground metro, where the main language is French. However, because they are dealing with the Egyptian government, they have to use English. They have to use the four skills at work.

Difficulties in four skills

He faces difficulty in reading when he doesn't know the meaning of words, he has to often refer to the dictionary. The main difficulty he faces is in writing, writing letters is a problem. As a knowledgeable engineer, he knows what he is talking about very well, but the transition from Arabic or from his studies, in order to explain takes a lot of effort and wastes about 60% of his time. He considers speaking as his second difficulty, listening as his third difficulty and finally reading. He is not looking for just one course but for an integrated course. He had previously started a course 2 years ago but it wasn't very suitable. He wanted speaking and he wanted a course that was faster. His teacher advised to go to a conversation course.

Description of "proficiency in English"

Fluency in speaking in complete sentences and words are clearly understood, not mixed up with something else. Clear pronunciation and does not have to clarify again. Perfect in all areas of English, not just speaking such as writing letters, discussion.

RECORD OF INTERVIEW 5: FEMALE, 18, LOWER INTERMEDIATE

Reason(s) for studying English

French is her first language since she studies in a French school (Notre Dame). She wants to improve her English because it is an importance language in this current era. She wants to have two languages at a good level because she wants to work in tourism or translation. She is a first year university student in the faculty of education, French section.

Use of English

She does not use English in her daily life but she needs English for her university because she takes it as a subject and she wants her English to be as good as her French.

Difficulties in four skills

When someone speaks English she cannot follow him easily. She understands the main ideas of what is being said but she cannot hold a discussion. She cannot speak with a foreigner and cannot speak at the same speed as Arabic. She cannot write a lot or with expression. She has limited ability and cannot express many ideas, however, she considers herself to be very good at grammar. She states that she is quite good at reading and can read simplified stories and that dictionaries help a lot. She plans to study English until she reaches a good level in English - about one year.

Description of "proficiency in English"

Pronunciation is very good, ability in grammar and writing is excellent. Any language needs to be used and language needs practice to learn new words. So a person at the top can lose this. Can concentrate very well.

RECORD OF INTERVIEW 6: FEMALE, 21, BASIC

Reason(s) for studying English

To improve her English because her level is very weak. She is studying English for her work in order to find a job. She was an art teacher and now she is engaged and her fiancé works in Oman where she'll need English to find a job.

Use of English

She does not use English in her daily life but at work in Oman she'll need since most of the employees over there are Indians and thus, English is necessary.

Difficulties in four skills

She can read words and write sentences that are dictated to her. However, pronunciation is difficult. In listening, she can understand some things. She considers her spoken level to be weak. She hopes to continue the courses until she reaches a good level in English in speaking. Work by its nature depends on speaking and no one will ask her about grammar.

Description of "proficiency in English"

By the way he speaks, his pronunciation, writing English. Can watch foreign serials without translation. Knows a lots of things in English, has memorized a lot of things.

RECORD OF INTERVIEW 7: MALE, 23, UPPER INTERMEDIATE

Reason(s) for studying English

He graduated from a language school (St. George) and went to the faculty of commerce for four years and forgot his English. He is studying English for computer studies and for the future.

Use of English

He works in a contracting company and does not currently use English.

Difficulties in four skills

Although he does not write much, he finds it difficult to express himself. He wants to improve his speaking but if he was stuck for a word, he would find another way to express himself. He can read, but if he did not understand a difficult word, he would try to guess it from the context of the sentence. He considers his listening to be good, depending on the accent. If it is American, he would find it difficult to understand, however the British accent is more familiar to us. He plans to finish all the courses and wants to speak English as fluently ("light tongue") and easily as he speaks Arabic.

Description of "proficiency in English"

Has no difficulty in speaking and doesn't search for words. Words are always available. Can express himself easily and has no difficulty. Does not have to think. Writing is the same thing. Can also read.

RECORD OF INTERVIEW 8: MALE, 23, LOWER INTERMEDIATE

Reason(s) for studying English

He wants to enroll in the Modern Accounting diploma at CACE. He plans to do a Masters degree abroad. He had previously enrolled for a postgraduate diploma in taxation at his university (faculty of commerce), however, it was all in Arabic and he did enjoy the professors way of teaching.

Use of English

He works but does not use English. He stated that in Egypt there is a big problem because English is used rarely at work. Even when you go to open an L/C at the bank, you don't need English. This is unlike the Gulf countries and Saudi Arabia where all those who work use English and no Arabic. All the information on the computers is in English. He added that he took computer course, and the teacher taught it in Arabic, but the expressions were in English.

Difficulties in four skills

He finds listening difficult because American speak very fast. He can listen to some words but not the whole sentence. He doesn't have time to keep up with a conversation. He is not used to speaking. His teacher had told him that his conversation was good, however, his opinion of himself is that he has a long way in front of him in order to speak with anyone. He has no difficulties in reading and writing and considers them good. He can write a paragraph and can read anything, even if there are difficult words to stop him. He wants to study English until the end because he eventually want to travel abroad to do a Masters degree.

Description of "proficiency in English"

He will be like an American. He has no difficulty in speaking. All the vocabulary will be available and will speak as he speaks Arabic. Regardless of the accent, he will speak easily.

RECORD OF INTERVIEW 9: MALE, 35+, BASIC

Reason(s) for studying English

He considers English to be essential for any job in Egypt or abroad. He was a marketing supervisor in Saudi Arabia and wants to continue to improve his English since any good position requires good English. He also applied for emigration to the US and is waiting for the result. He also stated that English is an essential language and gives a good personality and he hopes to teach his children in the future.

Use of English

He does not use English in Egypt as he has not worked here and has only been back for one month.

Difficulties in four skills

He finds listening to an American accent difficult. He can understand what is being said but some words are lost. He cannot say what he wants in English. He cannot respond with a convincing answer. In his writing, spelling is a problem. In Saudi Arabia, they appreciated that English was not his first language and focused on the content. He has no problems in reading and can read anything. He hopes the course would organize his knowledge and encourage him to speak without fear. He will continue as long as he finds himself improving.

Description of "proficiency in English"

Speaks well, reads well and can participate in any conversation in English. It is not someone who studied English the way we studied aiming only to pass from year to year.

RECORD OF INTERVIEW 10: FEMALE, 23, LOWER INTERMEDIATE

Reason(s) for studying English

The governmental schools do not produce a good level of English and students are unable to speak English at all. She wants to improve her language to have a better chance for a job. Looking at any newspaper ad, they all require English or computers.

Use of English

As a hostess on the Superjet she sometimes uses English with foreigners or tourists going to Hurghada, Luxor or Aswan. She added that English must be used and not only studied. We must speak it and listen to it.

Difficulties in four skills

She has difficulty in speaking English fluently. She can speak two sentences and then stops. She finds it difficult to discuss a topic. She can understand what is said but finds it difficult to respond. She knows the structure of a sentence: subject, verb, object. She can read any book, novel, dictionary. She tried to learn English on her own but couldn't. She plans to study until she reaches a level that satisfies her. She wants to be good and to be able to speak well - general English that she can use daily.

Description of "proficiency in English"

Was not able to give a description.

RECORD OF INTERVIEW 11: FEMALE, 18, UPPER INTERMEDIATE

Reason(s) for studying English

Now she is in the faculty of engineering and does not use English except for technical definitions. She feels she has forgotten almost all her English and is taking courses to maintain her level.

Use of English

She only uses English for her studies.

Difficulties in four skills

She finds it difficult to follow a conversation that is very fast. She can listen to songs on cassettes. She can speak with foreigners and friends but has difficulty in words. She has forgotten the pronunciation of many words and stated that there are many different accents and does not know which is correct. She has no difficulties in reading and writing. She hopes to be able to speak fluently.

Description of "proficiency in English"

Speaks English fluently almost like foreigners.

RECORD OF INTERVIEW 12: FEMALE, 23, INTERMEDIATE

Reason(s) for studying English

To improve her language because she is emigrating to Canada.

Use of English

Only uses technical language, expression in English. In the university her studies were mainly in Arabic.

Difficulties in four skills

She finds difficulty in understanding expressions and doesn't have time to understand the fast pace of words. She also finds difficulty in speaking. She considers her reading and writing to be reasonable as they depend on grammar and words. She is not used to speaking and listening.

Description of "proficiency in English"

Can use language in practical life well. Has perfect pronunciation and can read articles.

RECORD OF INTERVIEW 13: MALE, 35+, ADVANCED

Reason(s) for studying English

He works in a multi-national company and has some weak point in his English. He makes a lot of spelling mistakes and he wants to improve his position by perfecting his English.

Use of English

As a regional sales manager, he has to submit quotations and make weekly presentations. In his annual appraisal, his weakness in presentations and writing memos were indicated.

He studied at school and never took English courses. He studied business administration at university in Arabic.

Difficulties in four skills

As mentioned, he has difficulty in speaking and writing.

Description of "proficiency in English"

Perfect in the 4 skills.

RECORD OF INTERVIEW 14: MALE, 20, ELEMENTARY

Reason(s) for studying English

As a student in the faculty of arts, English section, Cairo university, he needs English. He got in by memorizing for the Thanawiya Amma (Egyptian high school certificate) and finds it difficult to understand English and needs private lessons.

Use of English

He does not use English in his daily life and only uses it for his studies.

Difficulties in four skills

He has difficulty in listening. He finds the tape fast and cannot follow the speed of lectures. He is shy and is afraid to speak to someone. He has spoken to the teacher but only a few words with foreigners. He has no problems in reading. In writing, he has problems in spelling, but overcomes that by memorization. He hopes to reach the required level (level 12) in order to study computers at CACE.

Description of "proficiency in English"

Someone who reached it by his effort. Very good in all the skills. Has good accent and speaks fast. Has good handwriting and spelling. Is confident in using the language.

RECORD OF INTERVIEW 15: MALE, 23-35, INTERMEDIATE

Reason(s) for studying English

He is a medical doctor and want to complete higher studies in the States. Therefore, he requires a score of 550 on the TOEFL.

Use of English

He deals with patients in Arabic, however, has to write the medical conditions and prescriptions in English. He uses English terminology and all the oral exams are in English.

Difficulties in four skills

In listening, can understand foreign serials but it is difficult to understand the American accent. The words in the street are not clearly understood. In speaking, has not dealt with foreigners. He plans to study until he passes the TOEFL.

Description of "proficiency in English"

Knows how to deal with others in a foreign country. Can handle conversations well with foreigners and can discuss specific points.

RECORD OF INTERVIEW 16: FEMALE, 23-35, ADVANCED

Reason(s) for studying English

Has finished medical school and has to pass a medical exam in the States which has a language component. She has to understand patients' complaints. She was in a French school and had been to the States before for a year.

Use of English

She does not use English since she has to speak in Arabic with her patients who hare mostly poor and uneducated. If English is used, it is British since a British syllabus is followed. American is totally different. Currently she lacks reading and does not have someone to talk to in English.

Difficulties in four skills

She is able to express herself in writing very well, however her vocabulary is not rich enough. She can get the general idea in reading but misses specific details in listening. During her one year in the States, she had a hard time understanding the teacher in class because she spoke so fast. She may have had good grammar, but it didn't help in the beginning. The lessons were based on hearing and watching film strips. She hopes to be reach the top and fill in the missing gaps.

Description of "proficiency in English"

It is not based on the pronunciation or accent. A person's fluency in language depends on vocabulary. Get the language from reading a lot and not from speaking with people. With rich vocabulary, good imagination one can write good essays. Good language is mainly vocabulary. To speak is to get used to the language from listening. She stated that Indians are good in language because they read a lot and have rich vocabulary but bad pronunciation.

Reason(s) for studying English

To find a good job. Has only a primary school certificate and is now studying second and third preparatory. Has tried many jobs. Some jobs do not need a certificate but definitely require language. She now works in a hotel and wants to go to a better hotel.

Use of English

For work with foreigners.

Difficulties in four skills

She has difficulty in speaking and does not understand what is said to her. When she took the exam she got very depressed because she knew nothing. She stated that she felt embarrassed to come and get her placement test results. She knows the alphabet and she can copy and form the letters. She can read some words. She wants to reach the stage where she doesn't speak a word of Arabic and has the words ready. She want to reach the same level as Arabic in all the skills.

Description of "proficiency in English"

Speaks very fluently. He has no hesitations and does not stop for words.

Reason(s) for studying English

She is currently a student in the faculty of commerce and all her studies are in Arabic. She is taking English in order to pursue computer studies and accounting in English at CACE. Because there are a large number of graduates from the faculty of commerce, knowing English would give her an advantage.

Use of English

She does not use English.

Difficulties in four skills

She considers her grammar to be quite good from school and stated that her work in commerce would not need grammar. She wants to improve her speaking. She stated that she was weak in writing compositions whether in Arabic or in English. She likes reading but hasn't tried to read a magazine or book in English. She said that she hasn't experience listening. When asked whether she finds watching English TV films and serials difficult, she responded that she does and has to read the translation.

Description of "proficiency in English"

Speaks English well, good pronunciation. Sentences are organized and meaningful. Clear way of speaking. Well-educated.

Reason(s) for studying English

He is a first year student in the faculty of commerce and stated that English is required for any job now. English is a must.

Use of English

Uses some English words with his friends when speaking Arabic. In one situation where he had to use English was during an interview when he applied for a job as a sales representative.

Difficulties in four skills

He finds listening to be a problem since he doesn't listen to a lot of English music. He got the TOEFL practice book and tapes in order to apply for AUC, but found he couldn't keep up with the tapes. He likes to speak English but pronounces some words wrongly. In reading, he has a similar problem and can't read some words. He considers his writing to be good and has no difficulties. He plans to finish the courses and then would like to sit for the CPA exam in his third year of university. He stated that while watching English films and focusing on the words then looking at the Arabic translation, has improved his vocabulary.

Description of "proficiency in English"

We can't say that he knows all the vocabulary, he knows a lot of words but occasionally may learn or benefit from new words. Speaks and writes well just like those on Nile TV. Someone like Hala Hashish who interviewed Al Gore on the "Good Morning Egypt" program. Way of speaking and pronouncing the words as Americans, not British. He prefers American.

Reason(s) for studying English

She is applying to a university in Canada and must take a TOEFL exam.

Use of English

She's a graduate of the faculty of arts, English department and teaches English in the university in Tanta. She uses English to some extent with the professors in the department, not in daily life. Occasionally with foreigners.

Difficulties in four skills

She finds listening to be very difficult, especially listening to a native speaker in different dialects. She does not find speaking English to be that difficult because if you have some vocabulary then you can speak. Writing and reading English are not difficult. The main problem is lack of practice. When speaking she cannot find the words spontaneously and she feels she has a problem with grammar. She stated that the secondary school syllabus is heavily loaded with grammar and there is no practice. She believes that living with native speakers is the way to improve one's language. She really wants to improve her English in order to travel abroad.

Description of "proficiency in English"

She admires people who speak English very well, their accent and likes the American accent. She stated that she speaks British but feels shy to speak to someone who is perfect in speaking with the same rhythm, not too fast and not too slow.

Reason(s) for studying English

Because of work and she hopes to travel to the US to visit her sister there. Also to help her children in the future.

Use of English

She doesn't use English except with friend who are also studying English courses.

Difficulties in four skills

She finds listening to be too fast because she can't catch some words but she can understand well. She considers her writing to be good. She reads well but there are some difficult words. She has some difficulty in grammar. She can understand some simple films. She plans to study until she can speak well and has no difficulty in speaking or understanding.

Description of "proficiency in English"

Speaks fast with no hesitations. Understands fast.

Reason(s) for studying English

He is a student in the faculty of arts, English department and finds studying English literature to be difficult because he needs to improve his English.

Use of English

Just for his studies and watches English films.

Difficulties in four skills

He finds speaking difficult, not easy to express himself. He finds speaking easier than listening. Listening is difficult, especially to foreigners. Writing is easier than speaking. He finds reading as difficult as listening. He can read specific sentence structures and write essays. He finds difficulty in reading novels and poetry, but newspapers are easier since normal words are used. He wants to improve his English to study other interest subjects such as hotels or tourism.

Description of "proficiency in English"

Has clear language and a clear point of view. Expresses his ideas simply and conveys meaning. Choice of words is appropriate. Can express himself easily.

RECORD OF INTERVIEW 23: FEMALE, 18-22, ELEMENTARY

Reason(s) for studying English

She wants to continue higher studies in sociology. English is essential for higher studies in reading references. Professors look at the English references in the bibliography. Many students are only able to read translations and not the originals. She doesn't want to find herself in such a situation.

Use of English

She stressed that English is the basis of everything and that someone who doesn't know English is ignorant, uneducated and not keeping up with the times. Pick up any toy or medicine and it is written in English. One feels handicapped at not being able to read it. English is a must.

Difficulties in four skills

Listening - she can understand everything the teacher says in the classroom. However, speaking is a problem. She can't form sentences or put sentences together. She has some problems in spelling. She prefers a written exam because to respond orally is very difficult. She feels like a child in spite of being an adult, she hesitates, is very shy. She can read a little and has some difficulty with words. She wants to know how to speak and deal with people in English.

Description of "proficiency in English"

Graduates of language schools, simultaneous translators. Their life is all English from KG. They speak English with each other, they don't speak Arabic. Their lives are in an English environment. They must read all the time or the language will be forgotten.

RECORD OF INTERVIEW 24: FEMALE, 35+, UPPER INTERMEDIATE

Reason(s) for studying English

She uses it in her job. She is a professor in the faculty of tourism, Ismailia. She is organizing a conference and the main language of the conference will be English. She will meet a lot of foreigners and has to speak with them.

Use of English

For work purposes. She hasn't taken any courses before. She has good French and Spanish.

Difficulties in four skills

She has difficulty in speaking because she doesn't use her vocabulary and grammar very well. She also needs to develop her listening. She has no problems in reading and writing. She wants to improve her speaking skills for the conference.

Description of "proficiency in English"

Fluent in conversation and 4 skills are developed, especially reading.

RECORD OF INTERVIEW 25:**FEMALE, 35+, INTERMEDIATE****Reason(s) for studying English**

She is sponsored by her job in the Ministry of International Cooperation and she uses English at work. All the agreements are written in English.

Use of English

For work purposes. She studied English from 1989-1992 then the funding stopped. The policy was changed once again, and so she is taking courses once again.

Difficulties in four skills

She finds difficulty in speaking. She went to the US via the Ministry and had to depend on colleagues to order food and help her in shopping. She found it difficult to listen to the American language because they eat up the words and speak fast. British is slower and easier. She has no problem in reading and writing. She plans to continue to study conversation until the end, funded by the Ministry. The only reason she stopped was that the Ministry stopped paying. English is an international language, used by all and it is better than any other language.

Description of "proficiency in English"

It is not by the speed of words, but the grammatical structure of the sentences. The basis of language is grammar. Has a large vocabulary, perfect the language. In writing, can put sentences together.

RECORD OF INTERVIEW 26: MALE, 18-22, BASIC**Reason(s) for studying English**

He is a student in the institute of tourism and hotels and needs English for his studies. He added that English is widespread.

Use of English

He sometimes uses English words and watches foreign films with translation.

Difficulties in four skills

He finds listening to be the most difficult. He can't transform English words to Arabic fast. He finds it difficult to listen to foreign songs. In reading, he can read at his own pace and translate into Arabic. He can write sentences. He has memorized some sentences / expressions: "How do you like our tour with us?" Speaking is easier than listening.

Description of "proficiency in English"

Fast and correct sentences with correct pronunciation without thinking, as if speaking Arabic. Like those who appear on the English TV broadcasts.

RECORD OF INTERVIEW 27: MALE, 18-22, LOWER INTERMEDIATE**Reason(s) for studying English**

He is a student in the faculty of commerce and to improve his language to study modern accounting at CACE.

Use of English

Does not use English for his studies. Uses a few words with friends.

Difficulties in four skills

Listening is most difficult, followed by speaking. Reading and writing are good. However, he is unable to read a magazine article on economics, there are many words he cannot understand. He can only understand 2 pages out of a 600-page book. He wants to study until he is accepted for modern accounting. He also plans to take conversation classes until the end.

Description of "proficiency in English"

Speaking and writing are excellent.

RECORD OF INTERVIEW 28: MALE, 35+, ELEMENTARY

Reason(s) for studying English

He needs it for his job in export and import which is all in English: figures, documents and certificates.

Use of English

Daily use at work. Deals with customers over the telephone. In previous job as an accounts auditor, he didn't need English.

Difficulties in four skills

Speaking is the most difficult. It is more difficult to speak to a person, it is easier on the telephone. When the person is in front of him he doesn't have any words and after the person leaves, he has the words. His listening is good and he can understand the meaning. Has no difficulty in reading. His handwriting is not good and uses a typewriter when required to write reports in English and colleagues help and correct his English.

Description of "proficiency in English"

Good way of speaking to others. Good accent and expressions use. Cannot distinguish whether Egyptian or American or British.

RECORD OF INTERVIEW 29: MALE, 23-35, ELEMENTARY

Reason(s) for studying English

He's an accountant and had worked for five years in Saudi Arabia where he did not use a single word of English. He now needs to practice his English because he tried to apply for several jobs. Computers will be introduced at his work and he may be sent abroad to study computerized accounting.

Use of English

He does not use English now. He watches a lot of video films. He stressed that English needs a lot of practice, otherwise it is forgotten.

Difficulties in four skills

His main problem is in speaking. He cannot speak and fears speaking with someone perfect. In a speaking situation he cannot speak, but after he leaves he keeps thinking he should have said this and that. He has a problem in writing. He considers his writing to be bad. He studies a lot of vocabulary.

Description of "proficiency in English"

Confidence in himself. Way of speaking with an excellent accent. Style of speaking, not hesitating.

RECORD OF INTERVIEW 30: FEMALE, 35+, ELEMENTARY

Reason(s) for studying English

She wants to learn English because she has a brother who lives in London and visits him and is thinking of living there in the future. She stated that English is very important and a person who does not know English is ignorant. It is no longer a luxury but it is essential. Now Egypt is trying to keep up with Europe and people from the lowest level should have a minimum level of English culture and speak English. We should also encourage tourism.

Use of English

She does not use English at home, only when she travels to London.

Difficulties in four skills

Speaking is most difficult for her. She can't arrange a sentence. Once when she was on the underground in London, she was speaking to a friend in English and someone made fun of her and she was very upset. She likes to watch foreign films and serials. She wants to study until she speaks and writes well and doesn't have to look at the translation when watching films.

Description of "proficiency in English"

From their accent. There is a difference between American and British English. The British pronounce English very well.

RECORD OF INTERVIEW 31: MALE, 35+, UPPER INTERMEDIATE

Reason(s) for studying English

He is an instructor in the air traffic control and academic studies institute and he felt that his English was starting to get worse. He was worried and so decided to improve his language.

Use of English

He uses English at work since aviation language is in English. He uses English in instructing the pilots, using specific phraseology. He studied English at school and has traveled a lot because he gets free tickets. He stated that his parents, father a professor at the agricultural college and mother at the faculty of science, Assiut university, used English at home.

Difficulties in four skills

He finds listening his weakest skill especially when listening to VOA or BBC. He has no problem in the other skills.

Description of "proficiency in English"

Fluent. It is very easy to understand him and it is very easy for him to express himself. It's easy for him to get out of a silly situation using the language.

RECORD OF INTERVIEW 32: MALE, 23-35, ADVANCED

Reason(s) for studying English

He is an air traffic controller and had a chance to study English at work and so he applied to join the course. He hopes for a good career opportunity in the International Civil Aviation Organization in Canada as a simultaneous interpreter.

Use of English

He uses English daily at work, using standard phraseology in order to avoid misunderstandings between pilots and air-traffic controllers. Some writing is involved such as flight registration, time-tables, ...

Difficulties in four skills

Finds difficulty in vocabulary and reading. His aim is to study simultaneous interpretation.

Description of "proficiency in English"

Good pronunciation and grammatical structure. Speaks correctly and reading and writing should also be good.

RECORD OF INTERVIEW 33: MALE, 18-22, ADVANCED

Reason(s) for studying English

He is a student at Ain Shams university, faculty of commerce. He was in school (not a language school) in Saudi Arabia where the level of English is higher than here. He is taking the course because he fears forgetting or losing the language. He has no opportunity to speak and the course will allow him to do so.

Use of English

He does not use English for his studies or in his personal life.

Difficulties in four skills

He find writing difficult because he cannot remember the spelling or structure. Right now writing is not important but will be when he works. He has no problem in other skills.

Description of "proficiency in English"

All the four skills are good. The method or style of writing, not spelling, is interesting and speaking is good.

RECORD OF INTERVIEW 34: FEMALE, 23-35, UPPER INTERMEDIATE

Reason(s) for studying English

She is a graduate of the faculty of commerce and now works as a secretary and accountant, office manager, in a foreign company. She works on the computer. She is taking the course to improve her English to go to a better job.

Use of English

She sometimes needs English in her job, mainly writing.

Difficulties in four skills

Grammar is one of her difficulties. Listening is also difficult. She can't get the meaning and can't get the words. She feels shy to speak English. When words are unfamiliar she has problems in reading and in writing.

Description of "proficiency in English"

Perfect. Knows when to say things. Produces correct sentences. Good in 4 skills. Good pronunciation and way of speaking. Know what he is saying.

RECORD OF INTERVIEW 35: FEMALE, 35+, BASIC

Reason(s) for studying English

She is a doctor, specializing in pediatrics and is doing her Ph.D. She also wants to apply to the AID program that runs the center for handicapped children for the Ministry of Health. Therefore, she needs English for her Ph.D. and possible travel.

Use of English

She uses English in reading and writing prescriptions. She reads more. She also needs English during medical professional development seminars.

Difficulties in four skills

Listening is the most difficult, especially in public. She has not had enough practice. She needs to listen twice in class. She may need a TOEFL course for traveling.

Description of "proficiency in English"

Can speak on every subject. Can listen and understand every subject and not just one's field. Can deal with ordinary people in an English-speaking community abroad.

RECORD OF INTERVIEW 36: FEMALE, 35+, LOWER INTERMEDIATE

Reason(s) for studying English

She is an air transport controller and therefore needs English for the job dealing with requests from foreign companies, embassies, negotiations with delegations.

Use of English

Needs English for the job and to attend conferences abroad. She studied English in the 1980s until level 18 in 1989 and stopped then. She has forgotten a lot and this is her first course.

Difficulties in four skills

She finds speaking difficult because she needs it with delegations. Quick responses are needed and to know what to say. AT the conference she found dealing with people to be difficult and the accents difficult to understand. They spoke very fast and she wanted to understand what they said.

Description of "proficiency in English"

Can deal with any situation in any place (eg. same level as those people in the delegations), can handle anyone. Reading and writing are essential. In reading looks at whole sentence and writes correctly.

APPENDIX 8:

TEACHERS' QUESTIONNAIRE QUALITATIVE DATA

Why do you think Egyptian adults study English? What are the different reasons that bring them to CACE?

1. To improve their English in order to find a good job. As a requirement so they can study computer or other studies. In order to be able to continue their studies in America or other countries.
2. For many young people, unemployment. These often go on to take business or computer courses. University students come to complement their academic studies once again in the hope of finding a good job. Graduate students come in the hope of studying abroad. Other CACE students are sent here by their companies. I suspect some students don't know why they're here.
3. Most of Egyptian adults study English to find good jobs or because it is needed at their work. Some young mothers study English to help their kids who go to language schools. Others are sent by their companies but these do not take it seriously and often come late and miss. They never do their homework.
4. Some of them study to master the language as it is an international language. Others need it to communicate with foreigners in their work in Egypt or abroad. Some need it to improve their cultural status to have a better chance of getting a good job. Others study to be able to continue their postgraduate studies or to join other departments in CACE as accounting or computer.
5. Cheaper than the British Council. I've heard to look for a husband or wife. Their company requires it. Need to study computer so need English Preparation to do TOEFL so can study further overseas. Help them to either get a job or improve their work situation. Some women specifically to help their children with English. Unemployed so filling in time improving themselves. Some, because they like English. Immigration overseas.
6. Egyptian adults study English because it's become a sort of gateway that they want to have wide open on the different fields of knowledge especially as related to economics, banking, modern technology, etc. CACE provides a relaxed learning atmosphere and good teachers.
7. Some students come to CACE to socialize. I think Egyptian adults study English because the rate of unemployment is increasing, thus some knowledge of English may give them privilege above other applicants for certain jobs or may help them be promoted to better posts. Besides, students need to learn English to be able to do other studies like business and computer studies. Students also like to receive certificates with the name of the American University.
8. Egyptian adults study English mainly for social aspects thinking that learning English gives them a sense of importance and individuality in the community, whether it be in the circle of family and friends or on the job market.
9. With the open door policy, Egyptian adults find it essential to study English as some of them are after a better position and others come to AUC to stimulate and extend their interest and skills in speaking and reading.
10. The reasons are: finding a suitable or a better job, studying computer science, for travelling or immigration, for cultural purposes, learning the language for 'language sake' as a necessity for social life.
11. They study English to improve their language. Some students need this language to get a job or a promotion or even a new post in another company. Others are obliged to take an English courses to be able to apply one of the diplomas or certificates offered in the CACE.
12. I think they study English because it's a must that they use it to get a decent job. Many reasons: the name CACE/AUC, they are sure of the material they get, the instructional aids used and how different the classes are from different institutes, the awareness of being instructed by qualified, professional instructors.
13. Mostly because they need it to get a good job. Some plan to travel or study (TOEFL). Fewer to help their children at school.
14. Egyptian adults are called upon to study English because of the intensive international contacts they have to cope up with namely politically, economically, scientifically and even across the media. Now, CACE in particular has an appeal among Egyptian over so many other institutions because of its fame, its established history and friendliness.
15. To improve their salaries, to get better respectable jobs, to help their children with their studies, to learn and work on computers.
16. To improve their career. To help their children at school. To be prepared for emigration sometimes. To apply for higher studies (MA) or other studies (computer).

17. Most of them didn't have a good chance to learn English when they were young. Besides, to them, English is easier to learn than any other language, you know, because of being familiar with American films, songs, etc. I think they come to CACE because of the good reputation, the location (Downtown, Heliopolis, Zamalek) and the fees which I think many people can afford.
18. Their main objective is a certificate. The certificate is used for work purposes. The certificate helps them in getting better jobs if they are at entry level. Once in a company English does not matter to them unless if they are competing with another person or move into export department or to further secure their jobs.
19. Work requirements, helping their children with school work, social identification.
20. Many come because of work, either to improve their language to get a good job or for promotion opportunities. Some companies encourage their employees to come and even pay for them. Others come for educational reason or because they're planning to emigrate or just for interest.
21. Maintaining their English. Better job prospects with a better knowledge of English. Socializing. A useful way of passing time.
22. To get a better job. To improve their language. To travel abroad. To be promoted in their work.
23. The students studying English literature or in the Faculty of Teachers join AUC to improve the speaking and listening skills. Students in others college hope to find good jobs or study ESP. Employees like to get better jobs or promotion. Accountants, bankers, tourist guides and physicians need it badly in their jobs.
24. It depends on the need they have. If they are high school students they just want to pass their final exams and /or get the highest grades to add to their grand total. If they are university undergraduates or graduates they realize the importance of learning English and they are self-motivated and achieve a lot.
25. For better job opportunities. To complete higher studies, MA/PhD. To socialize. For social prestige.
26. Job advertisements require the ability to speak English. Graduates who have to wait for years for a government job, doing nothing, take courses to kill the time. Some young ladies and divorced women think it's a good chance to make social relations. Employees who are sponsored by their companies may get a promotion, or at least get away from work for a few hours.
27. To be able to communicate with foreign people at work. To meet the needs of the place of work or study. To find a better job. To be able to travel abroad. To help their children with their schoolwork.
28. To go abroad. For work. AUC is prestigious and a certificate from AUC is appealing.
29. I think Egyptian adults come to CACE to improve their position at their jobs. Some of them are interested to travel or to immigrate. Most ladies who have children are eager to learn to be able to help their kids with their study, especially if they were in English schools.
30. I think that most adults study English for the following: to get a better job, to cope with the requirements of his present job, to travel abroad and communicate with other people, to make higher studies (reading for research work and writing thesis), to help children and kids with their assignments in language schools, to spend their sparetime in doing something useful and enjoyable.
31. They probably do because they need it for a better job opportunity. Seldom do they study it as a hobby. Sometimes it's a means to pursue their studies in the field of computer or business at CACE.
32. Better job opportunities. Career advancement. The possibility of further studies abroad. Enrolling in a graduate program. Self-improvement.
33. They are forced to come by their employers. they come dress to kill in the hope of gaining friends or even a mate. they would genuinely like to improve their English in the hope of getting better job opportunities. They want to travel and know that good English would facilitate their life abroad. They would like to boast that they're studying at AUC.
34. University students who want to better their chances of finding good jobs. Employees (drivers, cooks, secretaries) working with foreign companies. Mothers who have enrolled their kids in language schools. Employees sent by their work (paid for). Unemployed /waiting for the draft / a good way of spending time. PreCert. Students needing English to continue at CACE.
35. I think most of them come because they want to find a good job. Some of them want to travel abroad and few of them come for other reasons, such as helping their children at school or just because they love learning languages.
36. Students study English because they have better chances to improve their career. They need English to read references if they are making higher studies. It's very important for computer studies.
37. To enhance their social standard and their ability.

38. The main reason for studying English is to get a better job in either International companies or get promotions in their own jobs. The second good reason would be to upgrade their social standard especially that all language schools are private. As for CACE, it's AUC that bring them to us again to show off and meet people.

39. Improve school grades. Need it for work. Need it to help their children with their studies. Have free time on their hands, making friends. A pre-requisite for entering career certificates. To enhance their computer skills. Being able to read and understand written material for work. CACE reputation. AUC name. Fee is less when compared with British Council (may be more when compared with less reputable institutions). CACE teachers and teaching methods. Suitable days and hours. Meeting with friends and colleagues. Enjoying the facility with its widespread capabilities.

40. Most of them study English to get a better job at one of the joint-venture companies or banks. They also want to get a job at any of the resorts distributed throughout Egypt. Very few study it because they plan to emigrate.

41. Better job opportunities, class distinction, immigration, to qualify themselves for a promotion. It is now a prerequisite for university professors to get promoted so they should pass the TOEFL exam.

What are the different contexts or situations where your students would use English?

1. They would use English at work or when they travel to foreign countries.

2. In reality for most students the opportunities for using their English are extremely limited.

3. They use English at work in banks, at the airport, in foreign companies. Very few of them use English with their kids who need their help. Most of them never use the writing skill. That's why their writing is poor.

4. At work with foreigners. At home with their children. In the street to guide a tourist. Abroad to ask about accommodations. In an invitation.

5. This is a problem, many aren't in a situation to use it!! Working in oil / engineering / tourist companies where need to speak with foreigners, read and write letters / response. On computers. When immigrate. Have to present a thesis in English.

6. Most students use English in work, to study (especially postgraduate study) or to have a better chance for a better job. A few housewives learn English to teach their kids, but very few use it in everyday life.

7. They use them in secretarial and businesswork. Saleswork needs some English language too. Furthermore, those who work in the mass communication field need a certain level of English proficiency. For example those who work in journalism and public relations fields require the knowledge of English. Students sometimes need English to do postgraduate studies, make researches, translate articles, etc.

8. Mainly when speaking to a foreigner or travelling abroad to an English speaking country.

9. Doctors, engineers, programmers need to communicate with foreigners, to read references and to overcome the difficulty of being tongue tied.

10. These contexts or situations are: communication with others whether at home or abroad, interviews, job requirements (writing reports or making deals), conversation whether in office, at clubs, etc., talking with different nationalities that don't speak Arabic. English is the medium.

11. They usually use English at work; either with their boss or with foreigners and clients visiting them. Also, if they are studying or taking any course in the CACE (e.g. computer, hotels, ...etc.) they are obliged to use English because they use English books.

12. Through the past 2 years I noticed that students use English in making researches co-working with foreigners, traveling abroad for business or study, working in places related to tourism, communicating with foreign friends.

13. Speaking to foreign colleagues / bosses at work, writing memos.

14. Politics from simple to complex (you name it), business: businessmen/women, economics / banking, trade, secretarial / office work, scientific: doctors, engineers, lawyers.

15. To understand the language itself to work on computers, some work in hotels and restaurants. They need to know how to speak or respond to short questions.

16. If they are working for foreign organizations in Egypt. If they are involved in tourism in a way or another. If they have to give speeches or to lecture in English. If they are married to foreigners. If they have foreign friends. If they are required to attend lectures in English.

17. When you ask most of them why they study English they give you almost the same answer: work, business, travelling abroad.

- 18. At work especially in banks, secretarial positions, export field and those who want to learn computer studies, in the field of tourism and those who want to travel abroad for some reason or another.
- 19. Reading / writing formal business letters, sending / interpreting faxes, dealing with foreigners, teaching their children.
- 20. Many in work, dealing with foreign clients, business deals, airtraffic control, tourism. Many in studying various subjects. Also in watching foreign films, TV or reading foreign literature for pleasure.
- 21. Work, social.
- 22. At work. For studies. Travelling abroad.
- 23. Because of the privatization, university graduates use English in their jobs. Others deal with tourists. A few travel abroad. Parents like to help their kids study their lessons at home especially those who are in language schools. Few like to follow the English films and serials without looking at the Arabic translation.
- 24. Speaking with foreigners in places like embassies and foreign companies. Using English for correspondence with foreigners. Understanding and following up foreign movies and series.
- 25. In a work context and a study context.
- 26. Travel to another country. Talk to tourists. Work at a bank or a travel agency. Watch TV.
- 27. At work: to read reports, write required work, speak with foreigners,...etc. In social situations such as having a conversation with a tourist. Job interviews. Interviews at educational colleges and institutes.
- 28. At work. When speaking to a foreigner. When they are forced to do so in class.
- 29. Jobs, companies, dealing with foreigners in the street or at work if they are tourist guides.
- 30. At work (most banks and companies use English in their dealings with other offices). At home (helping kids and watching serials,...etc.). On a trip for fun or business or for study. For business and while attending conferences and meetings. For socializing with tourists and foreign visitors. For entertainment (watching TV programs, reading books and magazines in English and attending plays at theatres or seeing films at movies).
- 31. Some of the contexts are: using technical reports in English (or reading them), using English as a vehicle to study in other fields, in limited cases they need English for personal or direct communication with native speakers (i.e. people who work in international companies).
- 32. At work, especially where expatriates of foreigners are involved. Applying for a job. While travelling. Studies, e.g. computer course.
- 33. While studying abroad or even travelling for pleasure. Many jobs now require a good command of English. All jobs involving tourists and tourism require good English. All postgraduate studies require a certain knowledge of English.
- 34. Other than the classroom chances are nearly non-existent. If by chance they encounter English speaking foreigners or have pen friends.
- 35. At work. Writing to foreign pen friends. Abroad. Meeting foreigners in the streets of Egypt.
- 36. When they read references for their studies. When they meet foreigners. When they work for foreign agencies / companies.
- 37. Work and studies.
- 38. Their jobs whether International firms or Hotels where they deal with all different nationalities having English in common and travel agencies and tours again for tourists.
- 39. At work with colleagues or work associates. At home with the kids. With foreign friends visiting town. With foreign associates on business trips. Able to receive and convey messages through the phone (spoken) or fax / report (written).
- 40. With foreigners working at joint-venture companies and factories. With foreigners dealing with such companies. With tourists.
- 41. Classrooms, banks, tourist agencies, clubs, hotels, abroad (if they travel during the summer holiday).

Describe the language of a person at the top of the English proficiency continuum.

1. Could understand spoken and written English in various situations. Could be able to express himself in speaking and writing in a fluent way.
2. For me, this is someone who masters the language well enough to infer the meaning of phrases etc. which he hasn't learnt from their context. Someone for whom the language has come 'alive' and is no longer conscious of grammar / vocabulary (i.e. language tools) but expresses himself and understands automatically.
3. The written language of a person at the top of the English proficiency continuum includes a lot of spelling mistakes and a few grammatical ones. Orally he can express himself, understand and be understood but he must make mistakes of interference as well as grammatical ones.
4. He can communicate freely, effectively and fluently with others. He can form good expressive sentences which are correct grammatically and with advanced vocabulary.
5. Can read with expression and understanding novels, newspapers, magazines. Can hold heated discussions with good fluency. Listen without needing repetition by the speaker. They feel confident, comfortable in using English. Tell stories, jokes, understand a good range of English expression i.e. (kicked the bucket, she'll be apples mate, one for the road, ...) This general English - if a doctor specializing in a certain field, then would expect them to feel at home in the English of their field.
6. He should be able to communicate well (either in written or spoken language with native speakers and I say native because of the cultural differences that are usually manifest in the use of language).
7. He/she should be able to understand and communicate with a native speaker. He should be able to express him/herself and give his/her opinion in any topic whether verbal or written. Furthermore, he/she should be able to read English books in humanities and social sciences and make researches.
8. Sophisticated ideas, able to interpret idiomatic language. Able to discriminate technical and colloquial language. He is able to recognize compound words (words that are formed from more than one stem like 'Doubtfire') and words formed of initials. Able to use a dictionary and differentiate between the different categories mentioned in it.
9. The majority of advanced level students take the course seriously and they reach a very high standard of proficiency. The paragraphs they write in their final exam reflect their standard as they adhere to a great extent to the technique of paragraph writing.
10. That person should speak and write English perfectly. He should be able to express himself fluently and convincingly at any situation.
11. He should be able to express himself using suitable (correct) vocabulary and correct grammatical sentences fluently. .
12. At the top of the English proficiency continuum I think the student should be aware totally of his/her sentence structure, fluent in speaking near-native speakers in comprehension, 90% correct pronunciation, fully aware of language gestures, tones and intonation.
13. Fluent, clear accent, enough vocabulary to express oneself, correct grammar.
14. "Top" proficiency calls for an "absolute" ease at talking and thinking in English. It is mandatory that, accent left aside, proficiency implies fluency as much as promptness.
15. Masters the grammar. Good knowledge of vocabulary. Speaks easily and responds to any situation.
16. Fluent speaking, good expressive vocabulary and idioms in speaking and writing. High listening and reading comprehension. Enough awareness of cultural backgrounds.
17. It depends on your definition of the word "top", but I think it should be at least be able to communicate easily. He should have good command of the language, fluent able to read and understand English books, articles,...
18. I will describe within my knowledge of conversation 6. They can communicate in English well. Their vocabulary is less than satisfactory. Those with good vocabulary will tend to use the words not in their right context. For example, one student who's a biologist uses biology (medical) terminology when he's talking about a subject that doesn't deal with biology. The meaning is the same but it sounds strange. Their grammar is not always correct and when they get excited about a subject they always talk in Arabic to get their point across.
19. Someone who can use the language in most of the different situations like a native speaker with high fluency and accuracy.

20. At the top is someone who can communicate as a native speaker. Someone who can understand meanings of idioms from context.
21. They are false "advanced" learners with an intermediate level of English. Their passive knowledge of vocabulary is better than their production in English but their structure is often inaccurate.
22. Why not refer to books where answers are researched!!!!
23. Nearly as fluent as a native speaker. There should be no problem with the four skills in general.
24. S/He can easily express him/herself fluently either in speaking or in writing in any given situation. Through experience, s/he can build up accumulated data that could be of help throughout his/her life. It has been noticed that when somebody is excited and gets very nervous, s/he can easily resort to the vernacular unless s/he is fit enough to continue talking in English.
25. He/She should be fluent and be at ease using all 4 skills.
26. A person who can communicate with native and non-native speakers of English fluently and accurately.
27. It is very fluent and accurate language which needs minimum guidance and correction to achieve perfection quite easily.
28. Accurate and fluent provided the student is self-motivate and does a lot of reading, listening and speaking outside class.
29. To be able to understand English well. Be ready to communicate in different situations. Correct pronunciation is very important at this stage.
30. Could express and explain different points of view with the appropriate language free from grammar or spelling mistakes. Having the ability to use wide vocabulary and complex structures without unnecessary repetition. Correct pronunciation and mastering the 4 skills: listening, speaking, reading and writing. Fluency of production and creative thinking in English.
31. I think a person at the "top" of the continuum should attain a native-like proficiency regarding grammar, vocab, social and contextual language use and appropriateness.
32. This person should be fluent in speaking though the language may lack the natural use of idioms. There may be a few mistakes, some of which the speaker is aware of. Vocabulary is extensive and therefore authentic texts, e.g. newspapers and magazine articles and reports are easily read and understood. However, long literary texts may prove too arduous. Listening is no problem though not every single word may be grasped and there may be difficulty in understanding dialects, accents and slang. There will be very few mistakes in writing and these will tend to be stylistic and structural rather than grammatical.
33. He is fluent in both speaking and writing. He makes minimal mistakes while speaking. He is able to think in English which means he does not resort to translation either while speaking or writing. In my opinion, it would take a learner, with no knowledge of English, years to become proficient.
34. Unfortunately, there is no top. If adults come to CACE with inherent faults, those will often remain. However, if they start at the Beginners/Basic level they might be proficient. My idea of proficiency is someone who is not afraid to speak (fluent) and to communicate regardless of the grammar mistakes (accuracy) that may arise. However, I stress on adequate or near perfect pronunciation.
35. Very good writing skills (organization / developing ideas). Very good reading skills (skimming, scanning, etc.). Fluency in oral communication (ability to interact with English speaking persons with different accents, knowledge of formal and informal language). Good knowledge of English grammar and linguistics.
36. A person should be able to understand, read, speak and write the language correctly and fluently.
37. It should be rich and fluent. She/He should be able to talk fluently and think in English without any translation.
38. Can use the 4 skills fluently but most of the time keeps a certain accent especially in the sounds that are not found in our Arabic language "p", "b" and "th". Some could get closer to near native if they use their language frequently with non-native speakers of Arabic.
39. Excellent command of English with the ability of negotiating meaning and conveying self-expression and opinion in an extremely satisfying manner to the hearer / reader. Is able to maintain / sustain a steady flow of information with very minor pronunciation pitfalls that do not affect the flow.
40. S/He speaks fluently with a native-like accent, can discuss different topics in English and narrate past events. S/He can use the English lexicon properly. S/He is easily understood by native speakers of English. They don't have to ask him to repeat a word / a sentence or to clarify something that is not understood due to any kind of pragmatic failure that may take place due to the improper use of L2.
41. The learner should be able to articulate fluently and communicatively. His command of English structure and vocabulary should be ample. Makes effective use of tenses and lexis. No interference of the mother tongue. Should be able to communicate in formal and informal structure. Understands perfectly well complex structures and embedded sentences. Masters wide range of idiomatic use of English.

APPENDIX 9:
INSTRUCTIONS AND EXAMPLE ITEMS IN THE EPT
(ENGLISH LANGUAGE PLACEMENT TEST)

The American University in Cairo
CENTER FOR ADULT AND CONTINUING EDUCATION
Educational Assessment Unit

ENGLISH PROFICIENCY TEST

FORM

A 1

COPYRIGHT©1999

PLEASE DO NOT WRITE OR MAKE ANY MARKS IN THIS TEST BOOK

من فضلك لا تكتب أى شيء أو تضع أى علامة على كراسة الأسئلة

SECTION I: LISTENING COMPREHENSION

In this section of the test, you will have an opportunity to demonstrate your ability to understand spoken English. There are three parts in this section, with special directions for each part. Do **NOT** read ahead or turn the pages while the directions are being read. Do **NOT** take notes or write in your test booklet at any time.

Part (A) Statements

إختبار مهارة الاستماع الجزء (أ) الجمل الخبرية

فى هذا الجزء ستستمع إلى ١٥ (خمسة عشر) جملة خبرية. وحيث أن الجمل غير مكتوبة وستستمع لكل جملة مرة واحدة فقط فعليك بالتركيز لفهم ما يقوله المتحدث. بعد الاستماع لكل جملة انظر الى الصور الأربعة الموجودة فى كراسة الأسئلة ، ثم حدد أى صورة منها تعبر عن معنى الجملة التى استمعت إليها. على ورقة الإجابة وأمام رقم السؤال ظلل الدائرة التى تتفق مع إجابتك.

In Part A, you will listen to 15 statements. Each statement will be spoken just one time. The statement will not be written out for you. Therefore, you must listen carefully to understand what the speaker is saying. After you listen to a statement, look at the four possible pictures printed in your test booklet and decide which one expresses the closest meaning to the statement you have just heard. Then, on your answer sheet, find the number of the question and blacken the circle that corresponds to the answer you have chosen.

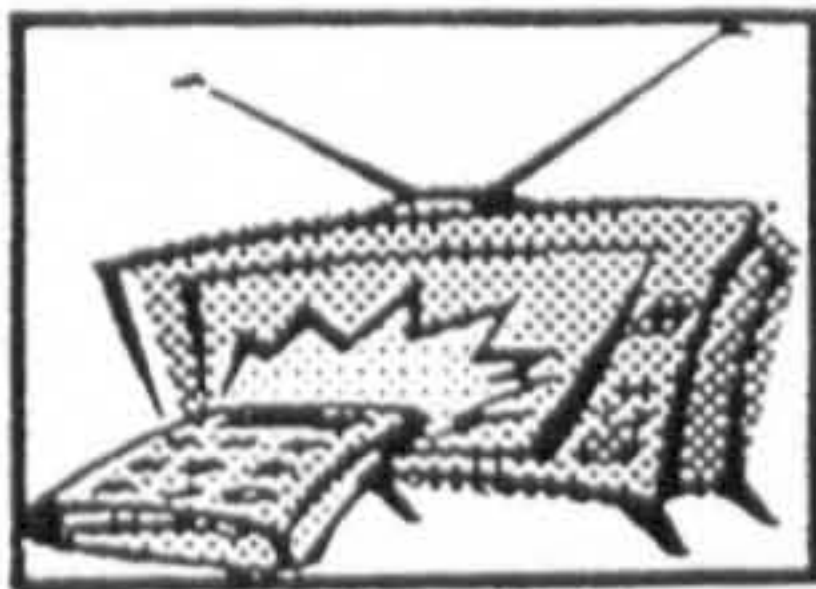
Now listen carefully to this example.

You will hear:

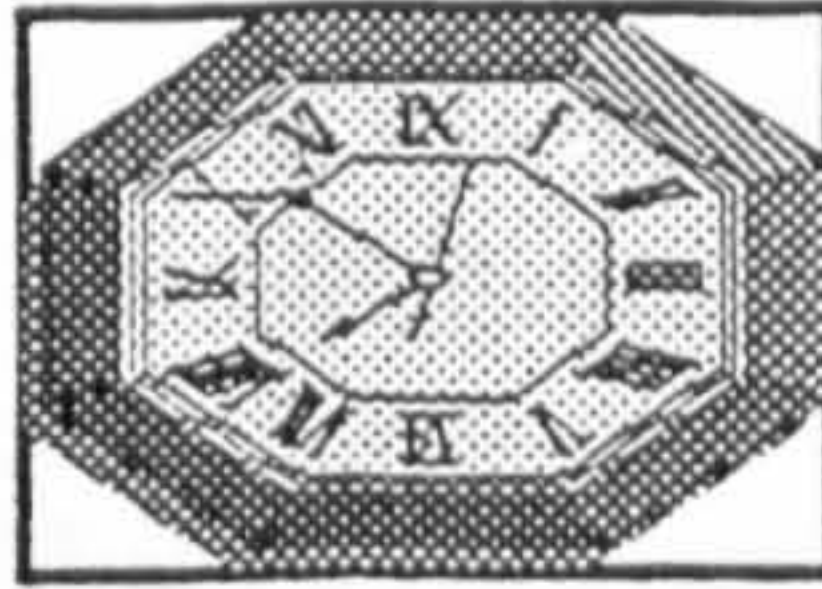
Sample Answer

(A) (B) (C) (D)

In your test booklet, you will have four pictures:



(A)



(B)



(C)



(D)

The speaker said, "This is my computer." Picture (C), "A computer" shows the same meaning of the statement. Therefore, you should choose answer (C). Note how the circle for letter (C) has been filled in on your sample answer.

Part (C) Mini-Dialogues

الجزء (ج) محادثات قصيرة

فى هذا الجزء ستستمع إلى ١٠ (عشرة) محادثات قصيرة. وفى نهاية كل محادثة سيقوم شخص ثالث بإلقاء سؤال عن المحادثة. اقرأ الاجابات الاربعة فى دفتر الاسئلة ، ثم حدد أفضل إجابة. وعلى ورقة الاجابة ظلل الدائرة المناسبة.

In Part C, you will listen to 10 short conversations between two speakers. At the end of each conversation, a third person will ask a question about what was said. You will hear each conversation and the question about it just **ONE** time. Therefore, you must listen carefully to understand what each speaker says. After you hear a conversation and the question about it, read the four possible answers in your test booklet and decide which **ONE** is the best answer to the question you have heard. Then, on your answer sheet, find the number of the question and blacken the circle that corresponds to the letter of the answer you have chosen.

Now listen carefully to this example.

You will hear:

Sample Answer

In your test booklet, you will read:

- (A) A pilot
- (B) An air hostess
- (C) A waitress
- (D) A security guard

(A) ☒ (C) (D)

From the conversation you learn that this dialogue took place during an air flight. The best answer to the question "Who is the woman?" is (B), "An air hostess." Therefore, you should choose answer (B).

Part (B) Question & Answer

الجزء (ب) سؤال وجواب

فى هذا الجزء ستستمع إلى ١٥ (خمسة عشر) سؤالاً. وحيث أنك ستستمع إلى كل سؤال مرة واحدة فقط، فعليك أن تستمع بتركيز. بعدها اختر من بين الإجابات الأربعة فى دفتر الأسئلة أفضل إجابة للسؤال الذى سمعته. وعلى ورقة الإجابة ظلل الدائرة التى تتفق مع إجابتك.

In Part B, you will listen to 15 questions. Each question will be spoken just one time. The questions will not be written out for you. Therefore, you must listen carefully to understand what the speaker is saying. After you listen to a question, read the four possible answers printed in your test booklet and decide which one would be the best answer to the question you have heard. Then, on your answer sheet, find the number of the question and blacken the circle that corresponds to the answer you have chosen.

Now listen carefully to this example.

You will hear:

Sample Answer

(A) (B) (C) ●

In your test booklet, you will read:

- (A) Half an hour ago.
- (B) From Cairo.
- (C) Platform No. 5.
- (D) In twenty minutes.

The speaker asked about the time the train leaves. The best answer to the question "When will the train to Luxor leave?" is (D), "In twenty minutes." Therefore, you should choose answer (D).

SECTION II: GRAMMAR
30 Questions - 20 minutes
(including the reading of the directions)

This section is designed to measure your ability to recognize language that is appropriate for standard English. There are TWO types of questions in this section, with special directions for each type.

Part (A): Filling in the blanks

الجزء الثاني
القواعد

التعليمات:

الأسئلة من ٤١-٥٥ هي جمل ناقصة وتحت كل جملة ستجد أربع كلمات أو جمل معرفة بالحروف (A), (B), (C), (D). اختر كلمة أو جملة واحدة تكون أنسب ما تكمل بها الجملة الناقصة. ثم أوجد رقم السؤال في ورقة الإجابة وظلل الدائرة المناسبة التي تتفق مع الحرف الذي يمثل الإجابة التي اخترتها.

Directions: Questions 41-55 are incomplete sentences. Beneath each sentence you will find four words or phrases, marked (A), (B), (C), and (D). Choose the ONE word or phrase that best completes the sentence. Then, on your answer sheet, find the number of the question and fill in the circle that corresponds to the letter of the answer you have chosen.

Example:

Sample Answer

Hurry! The train _____ in five minutes.
(A) left
(B) leaving
(C) has left
(D) is leaving

(A) (B) (C) (D)

The sentence should read, "Hurry! The train is leaving in five minutes." Therefore, you should choose answer (D).

Part (B) : Error Identification

التعليمات:

في الأسئلة من ٥٦-٧٠ كل عبارة لها أربعة كلمات أو جمل تحتها خط ومعرفة بالحروف (A), (B), (C), (D) وعليك ان تحدد كلمه او جملة واحدة ممن تحته خط والتي لابد من تغييرها لكي تصبح الجملة صحيحة. ثم أوجد رقم السؤال في ورقة الإجابة وظلل الدائرة المناسبة التي تتفق مع الإجابة التي اخترتها.

Directions: In questions 56-70, each sentence has four underlined words or phrases. The four underlined parts are marked (A), (B), (C), and (D). Identify the ONE underlined word or phrase that must be changed in order for the sentence to be correct. Then, on your answer sheet, find the number of the question and fill in the circle that corresponds to the letter of the answer you have chosen.

Example:

Sample Answer

The supermarket is on front of the station.
A B C D

(A) (B) (C) (D)

The sentence should read, "The supermarket is in front of the station." Therefore, you should choose answer (B).

SECTION III : READING COMPREHENSION

20 Questions

Total Time: 30 minutes

(including the reading of the directions)

الجزء الثالث

مهارة القراءة

ستقرأ فى هذا الجزء من الاختبار عديد من النصوص ويتبع كل نص بعض الأسئلة المبنية عليه ،
وعليك ان تختار فى الأسئلة من ٧١-٩٠ افضل إجابة (D) أو (C) , (B) , (A) طبقاً لما ورد فى
النص. ثم أوجد رقم السؤال فى ورقة الإجابة وظلل الدائرة المناسبة التى تتفق مع الإجابة التى
اخترتها.

In this section of the test, you will read a variety of reading texts. Each text is followed by
some questions about it. For questions 71-90, you are to choose the **ONE** best answer,
(A), (B), (C), or (D) to each question (according to the text). Then, on your answer sheet,
find the number of the question and fill in the circle that corresponds to the letter of the
answer you have chosen.

Now begin work on the questions.

THE AMERICAN UNIVERSITY IN CAIRO
CENTER FOR ADULT AND CONTINUING EDUCATION
EDUCATIONAL ASSESSMENT UNIT

Writing Placement Test

(30 Minutes)

You have been working in the accounting department for a “Petroleum Company” for two years now. You have not got any raise in your salary ever since.

Write a letter to your boss asking for an explanation for not receiving any increase during the past two years although you are a hardworking employee. You are also punctual and have good relations with everybody.

(Letter formatting is not part of evaluation of your writing)

الجامعة الامريكية بالقاهرة
مركز تعليم الكبار والتعليم المستمر
وحدة تقييم النشاط التعليمي

(٣٠ دقيقة)

اختبار تحديد مستوى القدرة على الكتابة

أنت تعمل في ادارة الحسابات لدى احدى شركات البترول لمدة عامين. ولكنك لم تحصل على اى زيادة في راتبك طوال هذه الفترة.

اكتب خطاباً لرئيسك طالباً تفسيراً لعدم حصولك على علاوة خلال العامين الماضين مع انك مجتهد ومنتظم في عملك ولك علاقات طيبة مع الجميع.

(الشكل العام لكتابة الخطاب ليس جزءاً من تقدير الدرجة)

ملحوظة:

- هذا الاختبار مترجم للغة العربية لسهولة الفهم فقط ولا بد ان تكون اجابتك كلها باللغة الانجليزية.

APPENDIX 10:

FOUR SAMPLE TRANSCRIPTS OF THE THINK ALOUD PROTOCOLS

Tapescript A

Male Elementary – 18 - 22

Writing

(The task is to write a letter. Some words are difficult. I am happy to find some Arabic. Reads prompt in Arabic. This is a translation of the prompt. It gives a better opportunity to answer and I will compare Arabic and English to get the meaning of the difficult words.)

(I wrote the date, I can't remember the rest. Samir and Ali. First I thank Mr. Samir and Ali.). I started the letter. This order is wrong. (meaning that you sent me the wrong order).

(Writing a letter is easier than writing a paragraph. I try to compare the English and the Arabic version of the prompt to know some words that I could use in writing the letter.)

He explains what he wrote. (I asked him to handle the problem. Then I sign.)

(The exam this way is suitable. The translation provides test takers with words they might not know. The time is also suitable, it is enough. the letter is easier than the paragraph as sometimes we do not know how to write sentences, but the letter is more structured. you write the date, first, then the words. even if one does not know how to write sentences, they may get points for writing the date,.. letter is always easy.)

Listening

3. b. (it says the young boy, then it is 'b'. it looks easy. Beautiful exam)

4. (the house with a garden, a' or it could be c')

5. d. (sleeping. it is a beautiful exam, like for children. The voice is not clear.)

6. c. it is 10:07 (hey. Be quick). *Sounds impatient at the slow pace.*

7. (it says fish) then 'd'.

8. I like my manager. 'a'. (Correct and wrong. What is this? I am not a child. It is a silly exam. aih da. Wala el atfal. Emtahan tafe. Is this KG 1 or KG2)

9. (a. also. Too slow. Why there is such a long pause. Come on. Yalleh ba'a)

10. d. the car the bus

11. (shirt. 'b')

12. b (because he says 28.)

13. a. number d. it is an easy exam.

(The first part of the exam is easy. It is as if for babies). OK. What is strange about this exam is that I listen and answer. It is new.

16. reads options. In Giza, of course.

17. *translates item into Arabic. Reads options.* 'c' (what am I doing?) I'm listening to the radio.

18. *translates item into Arabic. Reads options in English, translates '1'. 'b'.*

19. (All right.) *reads options*. No. inside the room. 'a'. no not inside the room.
20. *reads options*.
21. *reads only the correct answer aloud*. no c. you can borrow the books.
22. *reads options*.
23. *reads options, reads the correct choice*. I like tea without sugar.
24. *reads options*.
25. *reads options*.
26. *reads options*. [sound looks confident]
so far the exam is easy. The questions is OK, at the level of the student.
30. no 'd' *reads choice*. (Part a and part b are easy)
-

Part 'c'

31. *It stopped*.

(I finished a, and b'. I didn't do c' because it is beyond my level. The listening section in this new exam makes it difficult. That I listen and then choose the correct answer makes the student listens a lot and study more. And it take learning more seriously. It helps you understand conversation. It may be difficult at the beginning. In general, the listening is new and helps you focus on the exam. This is a very good way for the students.)

Grammar Section

Grammar is easy in general.

It is new that instructions are written in both Arabic and English.

41. *reads item and options*. There are. No. d'
42. *reads item and options*. Yes there is one. Yes there is one on the post office. No. b'
43. *reads item and options*. this is difficult. 'b'
44. *reads item and options*. Where no. 'b' (where)
45. *reads item and options*. Who repairs. No. b' also
46. *reads item and options*. No. d' questions are getting more difficult.
47. *reads item and options*. Will be have, no 'c'
48. *reads item and options*. Is used to processing. No. a computer is used for processing.
49. *reads item and options*. Should, 'b'
50. *reads item and options*. Might.
52. *reads item and options*. Very hot.

The exam staring from Section C, listening is getting more difficult and higher than my ability level.

I had difficulty understanding items 56 to 70. One needs wider range of vocabulary to be able to answer.

Instructions for this part were confusing.

The language used in these items is difficult. I am going to guess.

56. I guessed the answer and I am not sure of the answer.
58. *Reads and item and translates as he reads*. Guessing.
59. I also guessed the answer for this question.

60. Needs revision of grammar. Very difficult.

Items are new. Grammar is difficult. Vocabulary is also difficult.

Reading

71-73. It is getting more difficult. We are used to simple items. One should know that their ability is very low. Lower than that of KG1 or KG2. This exam requires us to listen, to read more and to put more effort in answering the exam. One is disappointed. Exams are really more difficult. One should know at least one or two words daily... the exam is very good as it makes us study more.

Reads next passage.

74-76

74. (This question is not direct. Needs more attention, the four options are similar. one needs a high level of language ability.

75. This question needs attention...

76. The word 'guarantee' I do not know how to pronounce it. I do not understand it and I can't use the context to guess its meaning.

Questions 77 – 80 are difficult and there are many difficult words. High concentration and it is hard for me. And I'll answer as far as I know or by guessing rather than the full understanding of the passage. I am not sure of the answer.

78. Is easy as it is in the passage. We read the passage and the options and get the correct answer.

79. The answer is in the passage.

80. It is hard to understand and I guessed the answer.

Questions 81-85 are very hard. When I get to a point in my reading, I forget the previous part. This needs more time. Time is not enough.)

Tapescript B

Female Elementary – 18 - 22

Writing.

You have been working in government a petroleum company. Yes. (easy, ok) for two years now. you have not got any raise in your salary ever since.

Write a letter to your boss asking for an explanation, (explanation ya'nee general).

(Although you are hard working) you are hard working employee.

(Some words are somewhat difficult but the topic seems easy. It is a good idea that the prompt is in Arabic, but it would have been better if one, in order to be self-dependent, we could, according to the level, provide translation of the difficult words only. This is better than translating the whole prompt. In this way, we will rely on ourselves.)

Reads prompt in Arabic.

(The idea of writing a letter may be a little difficult. I am still at the third level. I haven't learned how to write letters. I wrote on a general topic in my placement exam. I do not know what to write. I'll write what I know.)

Dear Manager

(I don't know what to write. Probably I am not used to this but I can say: certainly, this is going to be wrong, but I'll try. It is difficult in one more respect. Letter to my boss. this letter may have certain structure, certain organization. It could be easy as I will write about normal stuff. I have no problem with this but it is difficult in terms of structure and organization. I don't know how to write it.

There are words that I want to write in English, explaining the relationships between my friends and myself, and according to the letter, but I do not know these words because I haven't learned them yet. I am still in the third level. All I know is the past, how to form questions, I just know simple sentences that do not help me in expressing the ideas of the letter. The letter is easy though. but as I said, I'm still in the third level and there are words I want to use but I cannot because I do not know them.

The Arabic translation is good. there is only one thing wrong with it. I won't rely on myself to provide new words not mentioned in the translation. Only the difficult words should be translated into Arabic. Nothing more. It is also not appropriate. If a foreigner sees these materials in Arabic, they will think that we are making it too easy, as we don't know or we can't understand English. No. We need some challenge to learn to depend on ourselves and be able to use words on our own. This is similar to what happens in the prep and secondary school exams. We would use the same words in the prompt in the letter, changing the pronouns. This is because we did not know other words to use, and this is what I am doing now because I do not know any words to explain the relationship between the boss and his subordinates and how they want a raise. I hope there will be more English than Arabic so that we exert more effort.)

(I wrote something now which I am not sure it is right or wrong. I wrote:)

I want rise because life very expensive, now.

(I mean that I need a raise to be able to meet living expenses.)

It's wanted (it's required) everything now (I think it requires many things!. I want to write to cope with our times but this is very difficult for me to write. I don't know in English. but, we can say)

The world (I do not know the English for "develop")

The world ground every time (the world is growing!)

(or we can say I don't know the word for "develop") but I can say – write ground. (It means ('grow').

The world ground all time. (the world is developing continuously. I want to say it develops quickly.)

I can say 'all time'.(I am not sure if it is correct, but I can write it. it could be correct. It is my own words, that is why I want to use my own English. It could be wrong. No problem but I will learn. As they say, you need to work hard until you learn on your own.

These two sentences are completely my own. Maybe I won't take any credit for them, but they are my own. I can ask the teacher later.)

And I have two children (not two. This is not enough. I'll write) four children wanted (I have four children with many demands) more things. and the school. in (not 'and' lets'

say in) in school. (we can say) clothes and (vacation) all things (all these things) wanted
(require) raise the salary. (all these require a salary increase.)
That's enough. (to be sent to the director, asking for a raise and then sign.)

Listening

(This listening is very good.

I heard that if you want to learn a language, listen to its songs. Listening is good because
I try to imitate the sounds like the young baby who listens to their parents and imitating
them.

The example is 'c'.)

Young. number 'd'. (This is easy.)

My business card. My business card

(I did not hear. I was not paying attention.)

(I think it is 'b'.)

Number 'a'

(I'll choose number b)

Number c fishing food

number 'b' no no number d

(I do not understand number 9)

Number a

(There are hard words.)

Number a. (I can translate but I can't, I didn't hear.)

Number b, could be number c

Number a

Number d

Number c

(This is new to me. I haven't seen this part before.)

16. in Giza

17. Number c

18. Number 8, no number b

19. I did not hear

20. Number b

(Until now, some items are easy.)

21. (That was too fast.)

22. Number a.

23. Number b

24. (This is difficult. I didn't understand.

25. This is easy, but I don't understand.)

26. Number a. 'I don't

27. (I did not hear it.)

28. Number b

29. (That was difficult.)

(For a beginner like me, ten dialogues is too much.) *She sounded like: it's not fair.*
(It is too hard for me to answer ten spoken dialogues. Also the quality of the tape is not good).

31. Number a. (The tape quality is not good but it is a good idea.)

32. Could be number 'b'.

33. (That was difficult. I can't hear. Some words are spoken too fast for me.)

34. (That is above my level. I don't know the words. I am still at the beginner level.)

35. (He spoke too fast/ she said strange words.)

36. Number a. (I just guessed. It is too hard.)

37. (No, too hard.)

38. (I do not understand the questions. He is asking the question too fast. I do not know exactly what he is saying. I know that he is saying to her that he is meeting her tomorrow. she said what meeting. But I don't understand the question.) *Question posed after the dialogue.*

39. Could be number c. (I guessed.)

40. (No this is too difficult. I can not understand.) *She sounded like: save me!*

Grammar

(I know very little about grammar, as I graduated from a commercial school, not a high school)

Read instructions in Arabic. (I am familiar with these type of questions – multiple choice questions. Here is an example. The example item is easy. I hope that I find the rest of the items as easy as that. I am taking an English course to improve my grammar. I can speak well. I can read well. but it is hard to construct a sentence. Anyway. I will try.)

41. *Translates stem. Selects 'are'*

42. *Reads ticket office. Translates 'office'.* (I know how to translate. The items are easy if you study them. I haven't studied, so I write what I know. We can say 'on' the 'ticket office'). *She translated 'the ticket office' as 'desk' and so selects 'on'.*

(These two sentences are easy.)

43. *Misread the word 'ismailia' as ismaeel. Translated the rest of the sentence.* (I guess 'a'. By the way the sentences are still easy. I do not know the answer.)

44. *Translated correctly.* (Who for a person. where for a place. someone is asking where you buy the meat. or , how you do buy the meat? No. could be 'where'. maybe 'could'.)

45. *Misreads 'mechanic' as 'machine'.* (I don't know 45. I do not know the meaning of repair.)

46. *Mistranslates 'look the same' as 'looking at the same thing'.* Answer 'c'

47. Could be 'a'. *Translates.* (He is having a shower right now. I know that '-ing' is used to indicate continuity.)

48. *Reads item, with one option at a time.* I guess 'a'. I do not see him.

(Sentences are ok. I do not know some words, that is why I cannot answer.)

49. *Translates. Selects 'd', 'can use' wrongly assigning a passive meaning to it, as 'computer can be used'*

50. *Translates.* (So the answer is 'shouldn't', as he is very nervous.)
51. *Translates.* could be 'a'
52. *Translates.* the answer 'how'.
53. *Reads.* (The sentence is simple. I don't know some words.)
Proctor states that five minutes are left.
54. (There is not much time.)
55. *Translates.* (The cost.) *Translates* 'how many '. Then how much. for the cost.
-

Reads instructions. S(I have never done this type of questions.)

56. (This is really difficult.) *Reads.*
57. (What' shouldn't be written. It should be deleted.) answer 'a'
58. (Could be) 'c'
59. *Reads.* 'has' (is wrong.)
60. *Reads then translates.* (Sentences are tough. Grammar is hard. I am not used to this item type.)
61. Snsver 'a'
62. (I am still in level three. This is too hard. I am guessing.)
64.
(Time is over.)
-

Reading Comprehension

(This is a familiar format.)

Reads and translates. (Do not know 'announce'. I do not know 'schedule'.)

Misunderstands sentence. (I do not know 'skills'.)

(There are some difficult words. I will try to answer the questions.)

71. *Reads question, with long pauses.* (It is hard. The main idea may be easy. This is an announcement about training courses.)

72. (Number two is hard. I cannot translate it. Continues reading.

73. The meaning of the word 'conducted' is hard.)

Second passage

Reads the passage.

(This is a letter about summer vacation.

Misunderstands the information that the letter was sent to Mr. Ibrahim and so does not consider answers 'a' or 'b'. (Could be 'd'.)

75. Answer two nights.

(This is easy, but there are some difficult words. The underlined words are completely new to me. I cannot translate. I guess. I do not translate literally. I look at the overall meaning. I am not sure if it correct or wrong.)

76. Looks like 'promise'.

Passage three. *Reads.* (There are very difficult words. That is why I cannot answer the questions. Even at middle schools, comprehension passages were a lot easier.)

(There are heavy, difficult words. The comprehension section contains very difficult words and that is why I am not going to complete it.) *Test taker gave up.*

Tapescript G

Male – Intermediate – 23+

Writing

(This topic is completely new. The problem is that I haven't written any letters to companies before. Here I am writing to the GM. I want to convey to him that there was a mistake in the list of the company salary raise.) *Student did not understand prompt.* (Another problem is that I do not know some of the vocabulary. so here I am addressing a GM. I am not used to writing to a GM.) *Sounds a bit agitated.*

(I feel nervous because writing is difficult in itself. I lack the linguistic resources and so I have a problem in expressing the intended meaning. I'm experiencing uncertainty in choosing the right words. There is usually an introduction, for example, that one could easily use with any kind of letters such as writing a general letter to a friend. I am at a loss on what to write. A paragraph would have been better.

Another source of difficulty for me relates to the fact that I have never written about consequences. It is the most difficult part to write about consequences. In fact, one comes to the test worried.)

Students writes.

Listening section

(fast)

Part A

(Items require concentration. I have to concentrate to get the meaning. I think this might be very difficult for the beginner. After hearing the directions, I now understand what to do.)

1. (Time given for each answer is enough.)
2. (If one discovers a mistake, one does not have time to change it.)
3. (If one makes a mistake, there is a chance to make corrections.)
- 4.
- 5.
- 6.
- 7.
8. (Very good.)
9. (I did not have time. I did not understand the last part.)
- 10.
11. (I think this is for children at KG.)
12. (It is like children stories. Are we viewed as childish?)

(I was answering really fast as I was given a very short time to answer each item in this part, I relied on grammar. There are also some very light markings in the booklet, so this is probably the answer.)

Part B. (I think in order to save time, why not just read directions in Arabic only? I hope questions are easy so that I am able to answer all of them.)

- 13.
- 14.
- 15.
16. (There is time to read all the options.)
- 17.
18. (rekhem - Difficult question . This is better to assess level more accurately.)
19. (I did not hear.)
20. (I did not hear.)
21. (More difficult. I skipped three items. (God help me Istaghfar allah elazeem yarab)
- 22.
- 23.
- 24.
25. (I'm getting more and more tired of this exam.)
- 26.
27. (Thank god - elhamd lel lah, we are about to finish)
28. (All difficult.)
- 29.
- 30.

Part C

(Can't you say the listening items twice instead of one time, so that we hear them and can answer. Let's see.)

- 31.
32. (This is better.)
- 33.
- 34.
- 35.
36. (This part is better.)
- 37.
38. (Good question.)
39. (About to finish the listening.)
40. (I think there is a chance for cheating.)

Grammar section

41. *Reads item and options.* (I think there is. I didn't study grammar well.)

42. *Reads item and options.* I think 'next to'. (I am guessing. I think if one has reviewed grammar, one could answer all these items.)
43. *Reads item and options.* Where do you buy meet. 'b'.
- 44.
45. a mechanic is the man who.. repairs
46. *Reads item and options.* (Questions are good. These items are not grammar. They are 'special difficulties'.)
47. (Now this is grammar.) *Reads item and options.* (There aren't many grammar items.)
48. *Reads item and options.* 'b'
49. *Reads item and options.* You shouldn't. (It is confusing.)
50. *Reads item and options.* (In answering this section, one may select the correct form of the verb to complete that item and then if it is difficult one resorts to the meaning.
For example, in item '50', one resorts to the meaning. it could 'should' or 'shouldn't' I can't use grammar. moves to a new item. the meaning here is very important. I do not know the meaning of one word. I have studied this type of sentences and have heard them, so I know the correct answer once I read the stem.)
51. (Grammar questions are easy. This should be tenses. Like how to do things. These are easy questions.)
52. (Error identification is easy but sometimes (difficult - rehkem. God help us - Rabana youster)
53. The phrase, 'five-year plan'. (In this item it is the meaning that directs me in answering more than grammar. Here the question is not easy.)
54. *Reads item.* Many items.
- 55.
- 56.
- 57.
- 58.
59. *Reads item.* (What! I don't know this item.)
- 60.
61. (aah) (I thought errors in the words. Errors in grammar.)
- 62.
63. (Time is passing. Time allowed for answer is not written.)
64. f
65. f
66. d
67. d
68. d
69. d
- 70.

Part B is more difficult than the first. The sentence may be OK for me because I do not know the rule violated. Here there will be some hesitation. I feel nervous.

The mistakes in these sentences are simple and unnoticeable. it takes more time that the time allowed, 30 seconds.

Time was not enough. This makes me nervous. It also make me lose concentration. It makes me nervous when I read the item, to know that there is less than a minute makes me unfocussed. There is no concentration. I try to read the sentence. with difficult items, I read them over and over again, even though I know they are wrong. I feel that I spend too much time on them.

Reading

(Comprehension I don't know. Are we back to 'thanawaya ama'a'.) *Reads passage 71-73.*

(One reads the passage and understand it. I am reading the questions first.)

71. (There is a trick in question one. (ala min. I am very careful) items are tricky. One has to understand the comprehension well. Questions are not simple. Exam is long. There are many reading comprehension passages.)
72. (We are about to finish. Passage is getting more difficult. The last page. It looks too long. The most difficult. (bahareeze). Compares this to the old exam. The present exam is better in terms of item type. Items are not identical.)
- 73.

(I am asked about a word that I am not sure I know its meaning.)
Reads passage.

(The word is in the middle of the passage. I guessed it in the previous passage from context. I do not know the one in the present passage. The passage takes a long time to answer. Sometimes this underlined word is easy. Time is passing. This makes me look for the answer in the fastest way possible. The options are very similar. This makes it difficult. I feel that the reading passages get more difficult towards the end of this section.

I often left the question of what the writer means by this expression or this word. The reading comprehension is too long and difficult. There is more than one passage. One may make mistakes if passages are similar.)

Tapescript L

Female – Advanced – 23+

Writing

Reads writing prompt

(So you need a letter to my boss. This must be a memo. With my performance.)

Reads prompt again.

Subject:

Action:

Explanation:

(How can I start? Let's think of an appropriate introduction.)

Dear Sir:

As you know all of us are working to serve the AUC community. In addition, all of us are working to achieve one goal which is the highest performance. I am writing this memo to you to get your point of view of why I haven't got any increase during the past two years.

Reads letter (memo) again.

As you know I am one of your hard worker staff. I carry heavy workload in your department. Furthermore, I am punctual. Everyday, I start work at 8:30 a.m. and finish work at 3:30 p.m. on the other hand, if there is additional work, overtime without any remuneration because as you know that my work position does not allow me to get overtime paid. In addition, my communication skills are high. (Not this)- I have a good communication skills as same as interpersonal relations.

Listening comprehension

1. (This is very easy. Picture is clear.)
2. (There is a name and position. Clear too.)
3. (Dark classes give a hint to the picture.)
4. *Repeats item.* (This (picture) is not clear.)
5. (This is clear.)
- 6.
7. (The picture is not clear. Macdonalds makes fries and fish like the one in the picture.)
8. *Repeats item.* OK
9. *Repeats item.*
- 10.
- 11.
12. Easy
13. (Not clear. Maybe two getting acquainted. Maybe one is talking with another.)
14. *Repeats item.*
- 15.

Part 'B'

16. in giza
 17. A
 18. B. in the company's bus.
 19. On the second floor
 20. To catch the train
 21. A. take the next
 22. D
 23. Two and a half
 24. Sure. What are friends for
 25. It is terrific. (Three Cs in a row.)
 26. (This is not clear.) I think 'b'
 27. (I did not pay attention. I'll guess.) 'c'
 28. Eat less and exercise
 29. You know that I love reading. 'd'
 30. Take a break

 31. *Reads options.*
 32. repair shop
 33. *Reads options.*
 34. Mona is more helpful than Hala.
 35. d. (This is not clear. There are answers)
 36. he's not going.
 37. Refused
 38. C. she has not heard of tomorrow's meeting.
 39. (This is my strength.) Can use the computer well.
 40. Likes working and dealing with people.
-

Grammar section

40. *Reads item.* 'are' because 'restaurant.
41. *Reads item.* Next to
42. *Reads item.* Pleased. I guess
43. *Reads item.* Where?
44. *Reads item.* A mechanic
45. *Reads item.* Similar to each other, I guess
46. *Reads item.* He's having.
47. *Reads item.* I haven't seen him.
48. *Reads item.* Is used to. Is used for
49. *Reads item.* You shouldn't
50. *Reads item.* She might
51. *Reads item.* Very hot what is the weather like.
52. *Reads item.* It's going to build.
53. *Reads item.* There was no enough

54. *Reads item. Translates.* How much?
 55. *Reads item.* 'The' is wrong.
 56. *Reads item.* Where 's' is wrong.
 57. *Reads item.* 'would' I guess. A'
 58. *Reads item.* Was probably. 'b'
 59. *Reads item.* Wouldn't' is wrong.
 60. *Reads item.* You don't have the right... 'd' is wrong.
 61. *Reads item.* No. 'c'the father was taking care of the ...
 62. *Reads item.* 'c' these letters have to be mailed.
 63. *Reads item.* Twice. Because of. Not for
 64. *Reads item.* You will be completing
 65. *Reads item.* Couldn't have understood ...understand, understood, understood.
Couldn't have
 66. *Reads item.* Missed d
 67. *Reads item.* Is busy. 'is' is wrong.
 68. *Reads item.* The worst
 69. *Reads item.* Will be having, b
 - 70.
-

Section three reading comprehension

Reads passage.

71. *Reads item and options.* Reads answer. 'd'
72. *Reads item and options.* 'c'
73. *Reads item and options.* Held

Reads passage. Uses some Arabic. Translates portions of the reading as he reads.

74. *Reads item and options.*
75. *Reads item and options.* a.
76. *Reads item and options.* (I don't know. 'testify'.)

Reads passage. Uses Arabic again. (Select (five, hkamsa) applicants)

77. *Reads item and options.* filling in
78. *Reads item and options.* a chemistry. Here they did not mention
79. *Reads item and options.* 'c'
80. *Reads item and options.* to judge all the applicants or classify. Judge or classify. B.
classify. Student was hesitant.

Reads passage.

81. *Reads item and options.* He reads 'a'. No. 'b'
82. *Reads item and options.* Returns to passage. Does calculation in Arabic. D.

- 83. *Reads item and options. 'b'*
- 84. *Reads item and options. b'*
- 85. *Reads item and options.*

Reads passage.

- 86. *Reads item and options. how people, 'a'*
- 87. *Reads item and options. (This is difficult. I'll read it again.) Returns to passage. b.*
- 88. *Reads item and options. Pauses longer at the correct answer. 'c' reads answer*
- 89. *Reads item and options. quickly*
- 90. *Reads item and options. Increase*

Thank you.

APPENDIX 11:

FOUR SAMPLE RECORDS OF THINK ALOUD DATA ON ITEM RESPONSES

Test-Taker C: Male, Elementary Stage, 23+ age range

Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments
LA1	(√) It is easy.	LC31	()	GB61	(x)
LA2	(√)	LC32	()	GB62	(x)
LA3	(x)	LC33	(√)	GB63	(x)
LA4	(√) This is a good exam. It is comprehensive.	LC34	(x)	GB64	(√)
LA5	(√) Pauses between items are sufficient.	LC35	()	GB65	(√)
LA6	(√)	LC36	(√) Translates options.	GB66	(x)
LA7	(x)	LC37	(x) Difficult. I did not know.	GB67	(√)
LA8	(√)	LC38	() Tape is not clear at all.	GB68	(x)
LA9	(x)	LC39	(√)	GB69	(x)
LA10	(√)	LC40	(x) At the end items were difficult.	GB70	(x)
LA11	(√)		I am not good at grammar. God help me.	R71	(√) Difficult.
		GA41	(√)		
LA12	(x)	GA42	(x)	R72	(x) It is very difficult.
LA13	(√) Now the difficulty of the items increases. I am not sure of the last 2 items.	GA43	(x)	R73	(x)
LA14	(x) Really, it is getting more difficult.	GA44	(√)	R74	All this in 1 passage. God help me. (√) This passage is easier than last one.
LA15	(x)	GA45 -	(√) Reads items and options. Stops when he finds the correct answer.	R75	(√)
LB16	(√) Time is still sufficient to think of the answer. The exam is too long.	GA46	(x)	R76	(x)
LB17	(√)	GA47	(√)		Passages 1 and 2 took half an hour. Test taker gave up.
				R77	()
LB18	(√)	GA48	(√) Right now I really feel bored. The exam is getting too long.	R78	()
LB19	(√) Some of the items are still easy. Items are mixed in terms of difficulty.	GA49	(x)	R79	()
LB20	(√)	GA50	(x)	R80	()
LB21	(√)	GA51	(√)	R81	()
LB22	()	GA52	(x)	R82	()
LB23	(x)	GA53	(√)	R83	()
LB24	(√)	GA54	(√)	R84	()
LB25	(x)	GA55	(x) The last items are very difficult. I don't the meaning of some of the words.	R85	()
LB26	(x)	GB56	(√)	R86	()
LB27	(x)	GB57	(x)	R87	()
LB28	(√)	GB58	(x)	R88	()
LB29	(x)	GB59	(x)	R89	()
LB30	(√) My answers are not wrong. The tape is not clear.	GB60	(x)	R90	()

L: Listening
G: Grammar
R: Reading

A: Part A
B: Part B
C: Part C

Test-Taker F: Female, Intermediate Stage, 18 – 22 age range

Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments
LA1	(x) I heard 'young'	LC31	(x) Didn't hear.	GB61	(x) Reads item and translates. I don't know equipment. I am not sure. This is confusing.
LA2	(x) I didn't hear anything.	LC32	(√) Guessed.	GB62	(√)
LA3	(√) I could hear this one clearly.	LC33	(x) Time is not enough.	GB63	(x)
LA4	(x) I didn't hear this one.	LC34	(x) Not clear.	GB64	(√)
LA5	(√) I heard this.	LC35	(x)	GB65	(x)
LA6	(x) I heard this well.	LC36	(√)	GB66	(x)
LA7	(√) I guessed it.	LC37	(√) Can't hear.	GB67	(x)
LA8	(√) I guessed it too.	LC38	(x) Voice is not clear.	GB68	(√)
LA9	(x) I heard this..	LC39	(x) I don't know. I guessed a lot. There must be another chance.	GB69	(√)
LA10	(x) He should say the sentence twice.	LC40	(x) Didn't hear.	GB70	(x)
LA11	(√) I heard this. There are 2 options possible. He must repeat the sentence.	GA41	(√) This is a good item.	R71	Reading takes time. Half an hour is not enough. (x) It is hard. I do not know some words.
LA12	(x)	GA42	(√) Reads item.		(x)
LA13	(x)	GA43	(x) Very easy.	R72	(√)
LA14	(√)	GA44	(√).	R73	(√)
LA15	(√)	GA45	(√) Reads item and options. This is somewhat difficult.	R74	Now one has to read the questions first. (x) I don't know some words.
LB16	(√) Easy. I hope she or he reads it twice.	GA46	(x) Reads the item and options. Could be c or d.	R75	(√) Translated.
LB17	(√) The pause given between 2 items is sometimes sufficient and sometimes not.	GA47	(√) "He is having" because of 'now'.	R76	(x) I don't know.
LB18	(x) Not clear at all. There should be another reading.	GA48	(x)	R77	() I got really fed up, bored. I'll stop because it is boring. There are 5 passages and I have answered two.
LB19	(x)	GA49	(√) Translates.	R78	()
LB20	(√)	GA50	(√) Repeats correct answer	R79	()
LB21	(√)	GA51	(√) We can use 'will be' and 'going to' for the future. This way we get confused because both can be used. Good item.	R80	()
LB22	(x)	GA52	(√) Repeats correct answer	R81	()
LB23	(x)	GA53	(√)	R82	()
LB24	(x) I only heard a part.	GA54	(√)	R83	()
LB25	(x) Her voice is not clear. There is no chance to read the distracters of next item.	GA55	(√)	R84	()
LB26	(√) Her voice is not clear.	GB56	This is a confusing question. (x) Reads item. Revolution is a difficult word. There are 2 difficult words. I can't answer. Translates.	R85	()
LB27	(x) I guessed the answer.	GB57	(√) Reads item. Replaces 'what' with 'which'.	R86	()
LB28	(√) I don't know. There are 2 answers that fit.	GB58	(x) This is a silly question. I don't know.	R87	()
LB29	(x)	GB59	(√) Reads item.	R88	()
LB30	(√) His voice is clearer than hers. It should be said twice.	GB60	(√) Reads item.	R89	()
				R90	()

L: Listening
G: Grammar
R: Reading

A: Part A
B: Part B
C: Part C

Test Taker H: Female, Intermediate Stage, 23+ age range

Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments
LA1	(x) Good.	LC31	Some questions are too difficult. (x) Don't know.	GB61	() This is too hard and needs more time. I won't complete it. It's too difficult.
LA2	(x) Didn't hear.	LC32	(√) I heard some of this.	GB62	()
LA3	(x) I didn't understand.	LC33	(√) Reads options.	GB63	()
LA4	(√) Good. Some pictures are not clear.	LC34	(x) Reads options. Difficult.	GB64	()
LA5	(x)	LC35	(x)	GB65	()
LA6	(x)	LC36	(√) Reads options.	GB66	()
LA7	(x) Translated fishing	LC37	(x) I don't know.	GB67	()
LA8	(x)	LC38	(x) Reads options.	GB68	()
LA9	(x) Didn't understand. Need time to read.	LC39	(x)	GB69	()
LA10	(x) Talk, maybe telephone.	LC40	(x) I can't hear.	GB70	()
LA11	(x) Good. It would be difficult for the beginner.	GA41	(x) Reads item.	R71	(x) Reads items and options twice. Rereads passage.
LA12	(√)	GA42	(x) Reads item and options.	R72	(x) Reads items & options twice.
LA13	(√) Needs concentration.	GA43	(x) Reads item and options	R73	(√) Reads items & options twice.
LA14	(√)	GA44	(√) Reads item and options. Translates.	R74	(√) Reads items and options twice. Rereads parts of passage.
LA15	(x) Good. Very good.	GA45	(x) Reads item and options	R75	(x) Reads items and options twice.
LB16	(√)	GA46	(x) Reads item and options.	R76	(x) Reads items and options twice.
LB17	(x)	GA47	(x) Reads item & options.	R77	()
LB18	(√) Didn't hear. There are 2 correct answers.	GA48	(x) Reads item and options.	R78	()
LB19	(√)	GA49	(√) Reads item and options	R79	()
LB20	(x)	GA50	(√) Reads item & options. Translates	R80	()
LB21	(x) I guessed this one.	GA51	(√) This is an excellent part. If I couldn't answer some items, it is because I don't know and not because I can't hear.	R81	()
LB22	(x) Good.	GA52	(x) Reads item and options	R82	()
LB23	(√) If we depend on what we hear, it is difficult.	GA53	(x) Reads item and options	R83	()
LB24	(√) Getting difficult.	GA54	(√) Reads item and options	R84	()
LB25	(x)	GA55	(√) Reads item and options	R85	()
LB26	(x)	GB56	(x) I don't understand. Rereads directions.	R86	()
LB27	(x) Getting difficult.	GB57	(x) This part is difficult.	R87	()
LB28	(x) Don't know.	GB58	(x)	R88	()
LB29	(x) Didn't understand.	GB59	(√) More difficult. It needs good knowledge of grammar.	R89	()
LB30	(x) Reads options.	GB60	(x)	R90	()

L: Listening
G: Grammar
R: Reading

A: Part A
B: Part B
C: Part C

Test Taker I: Male, Advanced Stage, 18 – 22 age range

Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments	Item	Strategy(ies) / Comments
LA1	(√) Difficult for novice.	LC31	It is boring. I think MCQ is easier. An interview would be better. (√) It should be repeated twice. It is fast.	GB61	(√)
LA2	(√)	LC32	(x) It should be repeated twice. It is confusing.	GB62	(x)
LA3	(√) Could be fast for beginner levels.	LC33	(√) Some sentences are easy and some are difficult.	GB63	(x)
LA4	(√) Will be difficult for students up to Level 9. Should be repeated.	LC34	(x) I didn't hear that so my answer is wrong.	GB64	(√)
LA5	(√)	LC35	(√) Some sentences are easy and some are difficult.	GB65	(x)
LA6	(x)	LC36	(√) A lot of students will not be able to answer this type.	GB66	(√)
LA7	(√) This part will be difficult for many students	LC37	(x) Options are confusing.	GB67	(x)
LA8	(√) Picture might be confusing.	LC38	(√)	GB68	(x) .
LA9	(x) Should be repeated twice.	LC39	(√)	GB69	(x)
LA10	(√)	LC40	(√) Confusing.	GB70	(x)
LA11	(√) Even at the advanced level, one might not be able to hear some items.	GA41	(√)	R71	(x)
LA12	(√) Many people will be confused.	GA42	(√)	R72	(√) .
LA13	(x) 98 not 89.	GA43	(x)	R73	(√)
LA14	(√) Could be a, b or d.	GA44	(√)	R74	(√)
LA15	(x) Difficult for a lot of students. Very long.	GA45	(√)	R75	(√)
LB16	(√) Not easy. Should be repeated twice.	GA46	(√)	R76	(√)
LB17	(x) Very confusing. Not enough time to listen & select.	GA47	(√)	R77	(√)
LB18	(√) Concentrate.	GA48	(√)	R78	(√)
LB19	(√)	GA49	(x)	R79	(x)
LB20	(√) Listen carefully.	GA50	(√)	R80	(x)
LB21	(√) It might be too difficult for beginners.	GA51	(√)	R81	(x)
LB22	(x)	GA52	(x)	R82	(√)
LB23	(√) It should be repeated twice.	GA53	(√)	R83	(x)
LB24	(√)	GA54	(√) .	R84	(√)
LB25	(x) It is very confusing.	GA55	(√) .	R85	(x)
LB26	(√) Options are very confusing.	GB56	(x)	R86	(√)
LB27	(x) I couldn't hear.	GB57	(√)	R87	(√)
LB28	(x) Again.	GB58	(x)	R88	(x)
LB29	(x) Extremely difficult.	GB59	(√)	R89	(x)
LB30	(√)	GB60	(√)	R90	(x)

L: Listening
G: Grammar
R: Reading

A: Part A
B: Part B
C: Part C

APPENDIX 12:
MILES & HUBERMAN (1994:278-279) PROPOSED FRAMEWORK
FOR EVALUATION OF RESEARCH

OBJECTIVITY / CONFIRMABILITY

Relevant queries. Some useful questions to be asked:

1. Are the study's general methods and procedures described explicitly and in detail? Do we feel that we have a complete picture, including "backstage" information?
2. Can we follow the actual sequence of how data were collected, processed, condensed / transformed, and displayed for specific conclusion drawing?
3. Are the conclusions explicitly linked with exhibits of condensed / displayed data?
4. Is there a record of the study's methods and procedures, detailed enough to be followed as an "audit trail"?
5. Has the researcher been explicit and as self-aware as possible about personal assumptions, values and biases, affective states – and how they come into play during the study?
6. Were competing hypotheses or rival conclusions really considered? At what point in the study? Do other rival conclusions seem plausible?
7. Are study data retained and available for reanalysis by others?

RELIABILITY / DEPENDABILITY / AUDITABILITY

Relevant queries. What can be usefully asked in this domain?

1. Are the research questions clear, and are the features of the study design congruent with them?
2. Is the researcher's role and status within the study explicitly described?
3. Do findings show meaningful parallelism across data sources (informants, contexts, times)?
4. Are basic paradigms and analytic constructs clearly specified? (Reliability depends, in part, on its connectedness to theory.)
5. Were data collected across the full range of appropriate settings, times, respondents, and so on suggested by the research questions?
6. If multiple field-workers are involved, do they have comparable data collection protocols?
7. Were coding checks made, and did they show adequate agreement?
8. Were data quality checks made (e.g. for bias, deceit, informant knowledgeability)?
9. Do multiple observers' accounts converge, in instances, settings, or times when they might be expected to?
10. Were any forms of peer or colleague review in place?

INTERNAL VALIDITY / CREDIBILITY / AUTHENTICITY

Relevant queries. Some useful possibilities are:

1. How context-rich and meaningful (“thick”) are the descriptions?
2. Does the account “ring true”, make sense, seem convincing or plausible, enable a “vicarious presence” for the reader?
3. Is the account rendered a comprehensive one, respecting the configuration, and temporal arrangement of elements in the local context?
4. Did triangulation among complementary methods and data sources produce generally converging conclusions? If not, is there a coherent explanation for this?
5. Are the presented data well linked to the categories of prior or emerging theory? Do the measures reflect the constructs in play?
6. Are the findings internally coherent; are concepts systematically related?
7. Were rules used for confirmation of propositions, hypotheses, and so on made explicit?
8. Are areas of uncertainty identified? (There should be some.)
9. Was negative evidence sought for? Found? What happened then?
10. Have rival explanations been actively considered? What happened to them?
11. Have findings been replicated in other parts of the database than the one they arose from?
12. Were the conclusions considered to be accurate by original informants? If not, is there a coherent explanation for this?
13. Were any predictions made in the study, and how accurate were they?

EXTERNAL VALIDITY / TRANSFERABILITY / FITTINGNESS

Relevant queries. Here we may usefully ask:

1. Are the characteristics of the original sample of persons, settings, processes (etc.) fully described enough to permit adequate comparisons with other samples?
2. Does the report examine possible threats to generalizability? Have limiting effects of sample selection, the setting, history and constructs used been discussed?
3. Is the sampling theoretically diverse enough to encourage broader applicability?
4. Does the researcher define the scope and the boundaries of reasonable generalization from the study?
5. Do the findings include enough “thick description” for readers to assess the potential transferability, appropriateness for their own settings?
6. Does a range of readers report the findings to be consistent with their own experience?
7. Are the findings congruent with, connected to, or confirmatory of prior theory?
8. Are the processes and outcomes described in conclusions generic enough to be applicable in other settings, even ones of a different nature?
9. Is the transferable theory from the study made explicit?

10. Have narrative sequences (plots, histories, stories) been preserved unobscured? Has a general cross-case theory using the sequences been developed?
11. Does the report suggest settings where the findings could fruitfully be tested further?
12. Have the findings been replicated in other studies to assess their robustness? If not, could replication efforts be mounted easily?

APPENDIX 13:

FOUR SAMPLES OF THE COMPLETED TTSQ



الجامعة الأمريكية بالقاهرة
THE AMERICAN UNIVERSITY IN CAIRO
Office of Planning and Institutional Research (OPIR)
Testing and Evaluation Service (TES)

الرقم ٤٥
٤
٨٧
٤٥١
٤١٦٥
استقصاء حول
أساليب الإجابة على أسئلة الاختبار
عزيزي الدارس

يهدف الباحث من إجراء هذه الدراسة إلى التعرف على الأساليب التي اتبعها أثناء إجابتك على أسئلة الاختبار وكذا ريدود أفعالك تجاهه وسوف نستخدم المعلومات التي نحصل عليها من هذه الاستقصاءات في تخطيط وتصميم اختبارات تحديد المستوى اللغوي مستقبلاً بمركز تعليم الكبار والتعليم المستمر.

ونحن نتوجه إليك بالشكر على وقتك وتعاونك وبقدر إكمالك لهذا الاستقصاء حتى نهايته.

الاسم: د. محمد محمد محمد
رقم الدارس: ٤٥١٦٥

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعبه كل رقم بناءً على الشرح الموجود لكل رقم منها :

الرقم	المعنى / الشرح
١	لم استخدم هذا الأسلوب
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار
٣	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار

الرقم	المعنى / الشرح
١	١
٢	٢
٣	٣
٤	٤
٥	٥
٦	٦
٧	٧

115 KASR EL AINI STREET, P.O. BOX 2511 CAIRO, EGYPT TEL (202) 797-5964 FAX (202) 797-5967
426 FIFTH AVENUE, 3RD FLOOR NEW YORK, NY 10018-2729 TEL (212) 738-8800 FAX (212) 738-1600

الرقم	المعنى / الشرح
١	لم استخدم هذا الأسلوب
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار
٣	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار

الرقم	المعنى / الشرح
١	١
٢	٢
٣	٣
٤	٤
٥	٥
٦	٦
٧	٧
٨	٨
٩	٩
١٠	١٠
١١	١١
١٢	١٢
١٣	١٣
١٤	١٤
١٥	١٥
١٦	١٦
١٧	١٧
١٨	١٨
١٩	١٩
٢٠	٢٠
٢١	٢١
٢٢	٢٢
٢٣	٢٣
٢٤	٢٤
٢٥	٢٥
٢٦	٢٦
٢٧	٢٧
٢٨	٢٨
٢٩	٢٩
٣٠	٣٠
٣١	٣١
٣٢	٣٢

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما
يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها :

الرقم	المعنى / الشرح
٦	نوافق تماماً
٥	نوافق
٤	نوافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

الرقم	المعنى / الشرح
٦	نوافق تماماً
٥	نوافق
٤	نوافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

الرقم	المعنى / الشرح
٦	نوافق تماماً
٥	نوافق
٤	نوافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

الرقم	المعنى / الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب.

الرقم	المعنى / الشرح
٦	نوافق تماماً
٥	نوافق
٤	نوافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

الرقم	المعنى / الشرح
٦	نوافق تماماً
٥	نوافق
٤	نوافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

هل هناك أي أسلوب استخدمته أثناء إجابتك على هذا الاختبار لم ينكر في هذا الاستقصاء ؟
☐ نعم ☒ لا

إذا كانت الإجابة نعم من فضلك بين هذا الأسلوب وهل استخدمته مرة أو مرتين أم أكثر ؟

.....

.....

.....

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على التشرح الموضح لكل رقم منها

الرقم	المعنى / التشرح
٦	موافق تماماً
٥	موافق
٤	موافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

التفسير						العبارة
١	٢	٣	٤	٥	٦	
١	٢	٣	٤	٥	٦	٢١- شعرت بالإجهد وبدأت الإجابة عشوائياً بدون قراءة الأسئلة.
١	٢	٣	٤	٥	٦	٢٢- نسيت بعض الأشياء التي أعرفها بسبب التوتر.
١	٢	٣	٤	٥	٦	٢٣- كنت متأكدًا من الإجابة الصحيحة في أغلب الأسئلة.
١	٢	٣	٤	٥	٦	٢٤- شعرت أنه كان اختبار مشوقاً.
١	٢	٣	٤	٥	٦	٢٥- شعرت بالملل أثناء الاختبار.
١	٢	٣	٤	٥	٦	٢٦- كان الاختبار تجربة سارة.
١	٢	٣	٤	٥	٦	٢٧- أصابني الإحباط بسبب ضيق الوقت.
١	٢	٣	٤	٥	٦	٢٨- أعتقد أنها فكرة جيدة أن تكون التعليمات باللغتين العربية والإنجليزية.
١	٢	٣	٤	٥	٦	٢٩- أعتقد أنها فكرة جيدة أن يكون رأس السؤال في جزء الكتابة باللغتين العربية والإنجليزية.

• في كلمات مختصرة صف شعورك أثناء أدائك للاختبار في
 كلاً من: الجزء الأول الجزء الثاني الجزء الثالث الجزء الرابع الجزء الخامس الجزء السادس الجزء السابع الجزء الثامن الجزء التاسع الجزء العاشر الجزء الحادي عشر الجزء الثاني عشر الجزء الثالث عشر الجزء الرابع عشر الجزء الخامس عشر الجزء السادس عشر الجزء السابع عشر الجزء الثامن عشر الجزء التاسع عشر الجزء العشرون الجزء الحادي والعشرون الجزء الثاني والعشرون الجزء الثالث والعشرون الجزء الرابع والعشرون الجزء الخامس والعشرون الجزء السادس والعشرون الجزء السابع والعشرون الجزء الثامن والعشرون الجزء التاسع والعشرون الجزء الثلاثين

• في حالة شعورك بالتوتر أثناء الاختبار في أي جزء كان هذا الشعور؟
 في الجزء الأول الجزء الثاني الجزء الثالث الجزء الرابع الجزء الخامس الجزء السادس الجزء السابع الجزء الثامن الجزء التاسع الجزء العاشر الجزء الحادي عشر الجزء الثاني عشر الجزء الثالث عشر الجزء الرابع عشر الجزء الخامس عشر الجزء السادس عشر الجزء السابع عشر الجزء الثامن عشر الجزء التاسع عشر الجزء العشرون الجزء الحادي والعشرون الجزء الثاني والعشرون الجزء الثالث والعشرون الجزء الرابع والعشرون الجزء الخامس والعشرون الجزء السادس والعشرون الجزء السابع والعشرون الجزء الثامن والعشرون الجزء التاسع والعشرون الجزء الثلاثين

• هل اختلف هذا الشعور من جزء لآخر؟
 نعم لا



اساليب الإجابة على أسئلة الإختبار

يهدف الباحث من إجراء هذه الدراسة إلى التعرف على الأساليب التي اتبعتها أئمة إجماعتك على أسئلة الاختبار وكذا ردود أفعالك تجاهه وسوف نستخدم المعلومات التي نحصل عليها من هذه الاستقصاءات في تخطيط وتصميم اختبارات تحديد المستوى اللغوي مستقبلاً بمركز تعليم الكبار والتعليم المستمر.

الاسم : عبدالله مازة
رقم الدارس : 61106645

الرقم	المعنى/الشرح
٣	استخدمت هذا الأسلوب <u>عدة مرات</u> أثناء الاختبار.
٢	استخدمت هذا الأسلوب <u>مرة أو مرتين فقط</u> أثناء الاختبار.
١	<u>لم</u> استخدم هذا الأسلوب.

المتن			العبارة
١	٢	٣	
١	٢	٣	١- قمت بقراءة الأسئلة والاختيارات قبل أن أختار إحداهما.
١	٢	٣	٢- توقفت عن قراءة الاختيارات عندما توصلت للاختيار الذي بدأ لي صحيحاً.
١	٢	٣	٣- قمت بترجمة السؤال واختياراته.
١	٢	٣	٤- اخترت الإجابة الباقية بعد أن استبعدت ثلاثة اختيارات.
١	٢	٣	٥- جرب استخدام كل اختيار مع السؤال بشكل منفصل.
١	٢	٣	٦- لجأت للتخمين الذكي على أساس المعلومات المتاحة لي.
١	٢	٣	٧- حاولت الإجابة على السؤال قبل النظر إلى الاختيارات الواردة.

113 KASR EL AINI STREET, P O BOX 2511 CAIRO, EGYPT TEL (202) 797-5984 FAX (202) 797-5987
470 FIFTH AVENUE, 3RD FLOOR NEW YORK, NY 10018-2729 TEL (212) 750-8500 FAX (212) 750-1600

الرقم	المعنى / الشرح
٣	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب.

التمهيد			العبارة
١	٢	٣	
١	٢	③	٨- أعدت قراءة السؤال والاختيارات للاستيضاح.
①	٢	٣	٩- كنت أنجح الإجابة على السؤال أو اختيار الإجابة لوقت لاحق.
①	٢	٣	١٠- إذا كان هناك سؤالاً لم أهمه ، علم أنه له أية إجابة.
①	٢	٣	١١- قمت بعملية تجميع عشوائي.
١	②	٣	١٢- عبرت إجابتي عند اللزوم.
⑤	٢	٣	١٣- عندما لم أكن متأكد من الإجابة اخترت الإجابة الأطول أو الأقصر من بقية الإجابات.
⑤	٢	٣	١٤- عندما لم أكن متأكد من الإجابة سحنت عن الاختيار الذي كان يبدو مختلفاً عن الاختيارات الأخرى.
⑤	٢	٣	١٥- بعد الوقت مني قبل انتهائي من قراءة جميع الأسئلة.
⑤	٢	٣	١٦- حاولت الانتهاء من إجابة الاختبار بأسرع ما يمكن.
⑤	٢	٣	١٧- بدأت بقراءة سريعة للاختبار ككله.
⑤	٢	٣	١٨- كنت أراجع الوقت الذي انقضى والوقت المتبقى أثناء الإجابة.
⑤	٢	٣	١٩- كنت أقرب الآخرين لأرى مني يشبهون من الاختبار.
١	٢	③	٢٠- قرأت التعليمات جيداً.
⑤	٢	٣	٢١- تركت سؤالاً معييراً لإجابة.
١	⑤	٢	٢٢- قمت بمراجعة إجاباتي.
①	٢	٣	٢٣- توقعت طويلاً عند أحد الأسئلة.
			المسابيح المستخدمة للإجابة على جزء الكتابة :
١	٢	٣	٢٤- قرأت رأس السؤال باللغة الإنجليزية فقط.
١	٢	⑤	٢٥- قرأت رأس السؤال باللغة العربية - أولاً ثم بالإنجليزية.
١	٢	٣	٢٦- قرأت رأس السؤال باللغة الإنجليزية - أولاً ثم بالعربية.
١	②	٣	٢٧- قرأت رأس السؤال باللغة العربية فقط.
①	٢	٣	٢٨- أعدت قراءة رأس السؤال مرة أخرى.
⑤	٢	٣	٢٩- بدأت بصياغة أفكار بالغة بالعربية ثم ترجمتها إلى الإنجليزية وقمت بكتابتها.
①	٢	٣	٣٠- استخدمت في الإجابة على سؤال الكتابة كلمات من التي وردت في رأس السؤال.
١	٢	③	٣١- قمت بصياغة أفكار بالغة بالإنجليزية فقط.
⑤	٢	٣	٣٢- أعدت كتابة بعض الكلمات أو العبارات في إجابتي على سؤال الكتابة.

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رايك أمام كل عبارة من العبارات التالية طبقاً لما يبعثه كل رقم بناًما على الشرح الموضح لكل رقم منها :

المتغير						الملاحظة
١	٢	٣	٤	٥	٦	
١	٢	٣	٤	٥	⑥	١- بذلت قصاري جهدي في هذا الاختبار.
١	٢	٣	٤	٥	⑥	٢- أعتقد أنني أحسّت إجابة جيدة.
١	٢	٣	٤	⑤	٦	٣- شعرت أنه كان اختباراً سهلاً.
①	٢	٣	٤	٥	٦	٤- شعرت بالضغط للراحة الرغبة في حصولي على الإجابة من شخص آخر.
①	٢	٣	٤	٥	٦	٥- شعرت بأن الاختبار قد أصابني بالإرتباك.
①	٢	٣	٤	٥	٦	٦- شعرت أنه كان اختباراً صعباً.
١	٢	٣	٤	⑤	٦	٧- كنت مستعداً لهذا الاختبار.
١	٢	٣	٤	⑤	٦	٨- كنت أعرف ما ينبغي عليّ عمله أثناء الاختبار.
١	٢	٣	٤	٥	⑥	٩- أعتقد أنني كنت أعرف أهمية أن أبذل قصاري جهدي.
١	٢	٣	٤	٥	⑥	١٠- كان لدي الوقت الكافي للانتهاء من الاختبار.
١	٢	٣	٤	⑤	٦	١١- كان الاختبار بالنسبة لي تحدياً كبيراً.
١	٢	٣	٤	⑥	٦	١٢- شعرت بالإنجاز عندما أكملت الإجابة على الاختبار.
١	٢	٣	٤	⑤	٦	١٣- أحسست بالإرتباك عندما انتهيت من الاختبار.
①	٢	٣	٤	٥	٦	١٤- كنت متوتراً أثناء الاختبار.
①	٢	٣	٤	٥	٦	١٥- وجدت صعوبة في التركيز أثناء الاختبار.
①	٢	٣	٤	٥	٦	١٦- تحولت بصرى كثيراً في قاعة الاختبار أثناء الإجابة.
①	٢	٣	٤	٥	٦	١٧- أخشى الإختبارات عامة.
①	٢	٣	٤	٥	٦	١٨- أخشى الإختبارات لأنني أعرف أنها لا تظهر كفاشي كاملة.
١	٢	④	٤	٥	٦	١٩- شعرت بالإجهاد أثناء الاختبار.
①	٢	٣	٤	٥	٦	٢٠- لم أكمل الاختبار لأنه كان شديد الصعوبة.

هل هناك أي أسلوب استخدمته أثناء إجابتك على هذا الاختبار لم يذكر في هذا الاستقصاء ؟ ☒ نعم ☐ لا

إذا كانت الإجابة نعم من فضلكم بين هذا الأسلوب وأهل استخفتم مرة لو مرتين أم أكثر ؟
لم تعلم السنن وأضرباً عاماً من أسئلة الاستماع ولذا لم
كنت أجيب حسب ما استخفتم سماحة

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم باماً على الشرح الموجود لكل رقم منها .

الرقم	المعنى / الشرح
٦	وافق تماما
٥	وافق
٤	وافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

المتقدير						العبارة
١	٢	٣	٤	٥	٦	
①	٢	٣	٤	٥	٦	٢١- شعرت بالإجهاد وبدأت الإجابة عشوائياً بدون قراءة الأسئلة.
①	٢	٣	٤	٥	٦	٢٢- نسبت بعض الأشياء التي أعرفها بسبب تويزي.
١	٢	٣	④	٥	٦	٢٣- كنت متأكدًا من الإجابة الصحيحة في أغلب الأسئلة.
١	٢	٣	④	٥	٦	٢٤- شعرت أنه كان اختبار مشوقاً.
①	٢	٣	٤	٥	٦	٢٥- شعرت بالملل أثناء الاختبار.
١	٢	٣	٤	⑤	٦	٢٦- كان الاختبار تجربة سارة.
①	٢	٣	٤	٥	٦	٢٧- أصابني الإحباط بسبب ضيق الوقت.
١	٢	٣	٤	٥	⑥	٢٨- أعتقد أنها فكرة جيدة أن تكون التعليمات باللغتين العربية والإنجليزية.
١	٢	٣	٤	٥	⑥	٢٩- أعتقد أنها فكرة جيدة أن يكون رأس السؤال في جزء الكتابة باللغتين العربية والإنجليزية.

- [illegible]



الجامعة الأمريكية بالقاهرة
THE AMERICAN UNIVERSITY IN CAIRO
Office of Planning and Institutional Research (OPIR)
Testing and Evaluation Service (TES)

استقصاء حول أساليب الإجابة على أسئلة الاختبار

عزيزي الدارس

يهدف الباحث من إجراء هذه الدراسة إلى التعرف على الأساليب التي اتبعتها أثناء إجابتك على أسئلة الاختبار وكذا ردد أفعالك تجاهه وسوف نستخدم المعلومات التي نحصل عليها من هذه الاستقصاءات في تحليل وتصميم اختبارات تحديد المستوى اللغوي مستقبلاً لمركز تعليم الكبار والتعليم المستمر.

وبن نترجى إليك بالشكر على وقتك وتعاونك ومقدر إكمالك لهذا الاستقصاء حتى نهايته.

الاسم:
رقم الدارس: 672

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها :

الرقم	المعنى / الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب.

العبارة			التقدير
١	٢	٣	
١	٢	٣	١- قمت بقراءة الأسئلة والاختيارات قبل أن أختار إحداهما.
١	٢	٣	٢- توقفت عن قراءة الاختيارات عندما توصلت للاختيار الذي بدأ لي صحيحاً.
١	٢	٣	٣- قمت بترجمة السؤال واختياراته.
١	٢	٣	٤- إحترت الإجابة الباقية بعد أن استبعدت ثلاثة اختيارات.
١	٢	٣	٥- جرب استخدام كل اختيار مع السؤال بشكل منفصل.
١	٢	٣	٦- لجأت للتحمين الذهني على أساس المعلومات المتاحة لي.
١	٢	٣	٧- حاولت الإجابة على السؤال قبل النظر إلى الاختيارات الواردة.

113 KASR EL AINI STREET, P O BOX 2511 CAIRO, EGYPT TEL (202) 797-5984 FAX (202) 797-5987
426 FIFTH AVENUE, 3RD FLOOR NEW YORK, NY 10018-2729 TEL (212) 730-8500 FAX (212) 730-1600

الرقم	المعنى / الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب.

العبارة			التقدير
١	٢	٣	
١	٢	٣	٨- أعدت قراءة السؤال والاختيارات للاستيضاح.
١	٢	٣	٩- كنت أؤجل الإجابة على السؤال أو اختيار الإجابة لوقت لاحق.
١	٢	٣	١٠- إذا كان هناك سؤالاً لم أفهمه ، فلم أعط له أية إجابة.
١	٢	٣	١١- قمت بعملية تحمين عشوائي.
١	٢	٣	١٢- غيرت إجابتي عند اللزوم.
١	٢	٣	١٣- عندما لم أكن متأكدًا من الإجابة إحترت الإجابة الأطول أو الأقصر من بقية الإجابات.
١	٢	٣	١٤- عندما لم أكن متأكدًا من الإجابة بحثت عن الاختيار الذي كان يبدو مختلفًا عن الاختيارات الأخرى.
١	٢	٣	١٥- قمت بالوقت مسمي قبل انتهائي من قراءة جميع الأسئلة.
١	٢	٣	١٦- حاولت الانتهاء من إجابة الاختبار بأسرع ما يمكن.
١	٢	٣	١٧- بدأت بقراءة سريع للاختبار مأكلمه.
١	٢	٣	١٨- كنت أراجع الوقت الذي انقضى والوقت المتبقي أثناء الإجابة.
١	٢	٣	١٩- كنت أقرب الآخرين لأرى متى ينتهي من الاختبار.
١	٢	٣	٢٠- قرأت التعليمات جيداً.
١	٢	٣	٢١- تركت سؤالاً غير إجابة.
١	٢	٣	٢٢- قمت بمراجعة إجاباتي.
١	٢	٣	٢٣- توقفت طويلاً عند أحد الأسئلة.
			أساليب استخدمتها للإجابة على جزء الكتابة :
١	٢	٣	٢٤- قرأت رأس السؤال باللغة الإنجليزية فقط.
١	٢	٣	٢٥- قرأت رأس السؤال باللغة العربية - أولاً ثم بالإنجليزية.
١	٢	٣	٢٦- قرأت رأس السؤال باللغة الإنجليزية - أولاً ثم بالعربية.
١	٢	٣	٢٧- قرأت رأس السؤال باللغة العربية فقط.
١	٢	٣	٢٨- أعدت قراءة رأس السؤال مرة أخرى.
١	٢	٣	٢٩- بدأت بمصياغة أفكارتي باللغة العربية ثم ترجمتها إلى الإنجليزية وقمت بكتابتها.
١	٢	٣	٣٠- استخدمت في الإجابة على سؤال الكتابة كلمات من التي وردت في رأس السؤال.
١	٢	٣	٣١- قمت بمصياغة أفكارتي باللغة الإنجليزية فقط.
١	٢	٣	٣٢- أعدت كتابة بعض الكلمات أو العبارات في إجابتي على سؤال الكتابة.

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها

الرقم	المعنى / الشرح
٦	أوافق تماماً
٥	أوافق
٤	أوافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

العبارة						التفسير					
١	٢	٣	٤	٥	٦	١	٢	٣	٤	٥	٦
١	٢	٣	٤	٥	٦	٢١- شعرت بالإجتهاد وبدأت الإجابة عشوائياً بدون قراءة الأسئلة.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٢- نسبت بعض الأشياء التي أعرفها بسبب توترى.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٣- كنت متأكدًا من الإجابة الصحيحة في أغلب الأسئلة.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٤- شعرت أنه كان اختبار مشوقاً.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٥- شعرت بالملل أثناء الاختبار.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٦- كان الاختبار تجربة سارة.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٧- أصابني الإحباط بسبب ضيق الوقت.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٨- أعتقد أنها فكرة جيدة أن تكون التعليمات باللغتين العربية والإنجليزية.	١	٢	٣	٤	٥
١	٢	٣	٤	٥	٦	٢٩- أعتقد أنها فكرة جيدة أن يكون رأس السؤال في جزء الكتابة باللغتين العربية والإنجليزية.	١	٢	٣	٤	٥

• في كلمات مختصرة صف شعورك أثناء أدائك للاختبار ؟
Interesting and challenging

• في حالة شعورك بالتوتر أثناء الاختبار في أي جزء كان هذا الشعور ؟
When I was nervous

• هل اختلف هذا الشعور من جزء لآخر ؟
Yes, in the paragraph - the end of the test



المحكمة الأمريكية بالقاهرة
THE AMERICAN UNIVERSITY IN CAIRO
Office of Planning and Institutional Research (OPIR)
Testing and Evaluation Service (TES)

استقصاء حول أساليب الإجابة على أسئلة الاختبار

عزيزي الدارس

يهدف الباحث من إجراء هذه الدراسة إلى التعرف على الأساليب التي اتبعها أثناء إجابتك على أسئلة الاختبار وكذا ردود أفعالك تجاهه وسوف نستخدم المعلومات التي نحصل عليها من هذه الاستقصاءات في تخطيط وتصميم اختبارات تحديد المستوى للوعي مستقبلاً بمركز تعليم الكبار والتعليم المستمر.

ونحن نتوجه إليك بالشكر على وقتك وتعاونك ونقدر إكمالك لهذا الاستقصاء حتى ننتهيه.

الاسم: **الشيخ التوفيق خير**
رقم الدارس: **3404339**

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها.

الرقم	المعنى / الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب

الترتيب	العبارة
١	١ - قمت بقراءة الأسئلة والاختيارات قبل أن أختار إحداهما.
٢	٢ - توقعت عن قراءة الاختيارات عندما توصلت للاختيار الذي بدأ لي صميحاً.
٣	٣ - قمت بترجمة السؤال واختياراته.
٤	٤ - اخترت الإجابة الباقية بعد أن استبعدت ثلاثة اختيارات.
٥	٥ - جرب استخدام كل اختيار مع السؤال بشكل منفصل.
٦	٦ - لجأت للتخمين النكبي على أساس المعلومات المتاحة لي.
٧	٧ - حاولت الإجابة على السؤال قبل النظر إلى الاختيارات الواردة.

113 KASR EL AINI STREET, P.O. BOX 2511 CAIRO, EGYPT TEL (202) 797-5984 FAX (202) 797-5987
420 FIFTH AVENUE, 3RD FLOOR NEW YORK, NY 10018-2729 TEL (212) 730-3500 FAX (212) 730-1600

الرقم	المعنى / الشرح
٢	استخدمت هذا الأسلوب عدة مرات أثناء الاختبار.
٢	استخدمت هذا الأسلوب مرة أو مرتين فقط أثناء الاختبار.
١	لم استخدم هذا الأسلوب

الترتيب	العبارة
١	٨ - أعدت قراءة السؤال والاختيارات للاستيضاح.
٢	٩ - كنت أؤجل الإجابة على السؤال أو اختيار الإجابة لوقت لاحق.
٣	١٠ - إذا كان هناك سؤالاً لم أهتم به، فلم أعط له أية إجابة.
٤	١١ - قمت بعملية تخمين عشوائي.
٥	١٢ - عبرت إجابتي عند اللزوم.
٦	١٣ - عندما لم أكن متأكدًا من الإجابة اخترت الإجابة الأطول أو الأقصر من بقية الإجابات.
٧	١٤ - عندما لم أكن متأكدًا من الإجابة بحثت عن الاختيار الذي كان يبدو مختلفاً عن الاختيارات الأخرى.
٨	١٥ - بعد الوقت مضي على امتحاني من قراءة جميع الأسئلة.
٩	١٦ - حاولت الانتهاء من إجابة الاختبار بأسرع ما يمكن.
١٠	١٧ - بدأت بقراءة سريعة للاختبار مأكلمه.
١١	١٨ - كنت أراجع الوقت الذي انقضى والوقت المتبقي أثناء الإجابة.
١٢	١٩ - كنت أقرب الآخرين لأرى متى ينتهون من الإختبار.
١٣	٢٠ - قرأت التعليمات جيداً.
١٤	٢١ - تركت سؤالاً بغير إجابة.
١٥	٢٢ - قمت بمراجعة إجاباتي.
١٦	٢٣ - توقفت طويلاً عند أحد الأسئلة.
١٧	أساليب استخدمتها للإجابة على جزء الكتابة:
١٨	٢٤ - قرأت رأس السؤال باللغة الإنجليزية فقط.
١٩	٢٥ - قرأت رأس السؤال باللغة العربية - فولا تم بالإنجليزية.
٢٠	٢٦ - قرأت رأس السؤال باللغة الإنجليزية - فولا تم بالعربية.
٢١	٢٧ - قرأت رأس السؤال باللغة العربية فقط.
٢٢	٢٨ - أعدت قراءة رأس السؤال مرة أخرى.
٢٣	٢٩ - بدأت بصياغة أفكارى باللغة العربية ثم ترجمتها إلى الإنجليزية وقمت بكتابتها.
٢٤	٣٠ - استخدمت في الإجابة على سؤال الكتابة كلمات من التي وردت في رأس السؤال.
٢٥	٣١ - قمت بصياغة أفكارى باللغة الإنجليزية فقط.
٢٦	٣٢ - أعدت كتابة بعض الكلمات أو العبارات في إجابتي على سؤال الكتابة.

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها :

الرقم	المعنى / الشرح
٦	نوافق تماماً
٥	نوافق
٤	نوافق إلى حد ما
٣	نغير موافق إلى حد ما
٢	نغير موافق
١	نغير موافق على الامتلاك

المتغير						المسألة
١	٢	٣	٤	٥	٦	
١	٢	٣	٤	٥	①	١- بدت قصاري جهدي في هذا الاختبار.
١	⑥	٣	٤	٥	٦	٢- أعتقد أنني أحت إحالة حيدة.
١	٢	③	٤	٥	٦	٣- شعرت أنه كان اختبارا سهلا.
①	٢	٣	٤	٥	٦	٤- شعرت بالصعب لدرجة الرغبة في حصولي على الإحالة من شخص آخر.
١	٢	٣	④	٥	٦	٥- شعرت بأن الإختبار قد أصابني بالإرتباك.
١	٢	٣	⑤	٥	٦	٦- شعرت أنه كان اختبارا صعباً.
١	٢	٣	٤	⑥	٦	٧- كنت مستعداً لهذا الاختبار.
①	٢	٣	٤	٥	٦	٨- كنت أعرف ما ينبغي عليّ عمله أثناء الإختبار.
١	٢	٣	٤	⑥	٦	٩- أعتقد أنني كنت أعرف أهمية أن أبذل قصارى جهدي.
١	٢	٣	⑤	٥	٦	١٠- كان لدى الوقت الكافي للانتهاء من الاختبار.
١	٢	٣	٤	⑥	٦	١١- كان الإختبار بالنسبة لي تحدياً كبيراً.
١	٢	٣	⑥	٥	٦	١٢- شعرت بالإنجاز عندما أكملت الإجابة على الاختبار.
١	٢	٣	⑤	٥	٦	١٣- أحسست بالإرتياح عندما انتهيت من الاختبار.
①	٢	٣	٤	٥	٦	١٤- كنت متوتراً أثناء الاختبار.
١	٢	٣	⑤	٥	٦	١٥- وجدت صعوبة في التركيز أثناء الاختبار.
①	٢	٣	٤	٥	٦	١٦- تحولت بعصري كثيراً في قاعة الاختبار أثناء الإجابة.
١	٢	٣	٤	⑥	٦	١٧- أخشى الإختبارات عامة.
١	٢	٣	④	٥	٦	١٨- أخشى الإختبارات لأنني أعرف أنها لا تظهر كفاشي كاملة.
١	٢	٣	⑤	٥	٦	١٩- شعرت بالإجهد أثناء الاختبار.
١	٢	٣	①	٥	٦	٢٠- لم أكمل الاختبار لأنه كان شديد الصعوبة.

1

يرجى التفضل برسم دائرة حول الرقم الذي يعبر عن رأيك أمام كل عبارة من العبارات التالية طبقاً لما يعنيه كل رقم بناءً على الشرح الموجود لكل رقم منها:

الرقم	المعنى / الشرح
٦	موافق تماماً
٥	موافق
٤	موافق إلى حد ما
٣	غير موافق إلى حد ما
٢	غير موافق
١	غير موافق على الإطلاق

العبارة						التعليق
١	٢	٣	٤	٥	٦	
١	٢	٣	٤	٥	٦	٢١- شعرت بالإحباط وبدأت الإحالة عشوائياً بدون قراءة الأسئلة.
١	٢	٣	٤	٥	٦	٢٢- نسبت بعض الأشياء التي أعرفها بسبب توترى.
١	٢	٣	٤	٥	٦	٢٣- كنت متأكداً من الإجابة الصحيحة في أغلب الأسئلة.
١	٢	٣	٤	٥	٦	٢٤- شعرت أنه كان اختبار مشوقاً.
١	٢	٣	٤	٥	٦	٢٥- شعرت بالملل أثناء الاختبار.
١	٢	٣	٤	٥	٦	٢٦- كان الاختبار تجربة سارة.
١	٢	٣	٤	٥	٦	٢٧- أصابني الإحباط بسبب ضيق الوقت.
١	٢	٣	٤	٥	٦	٢٨- أعتقد أنها فكرة جيدة أن تكون التعليقات باللغتين العربية والإنجليزية.
١	٢	٣	٤	٥	٦	٢٩- أعتقد أنها فكرة جيدة أن يكون رأس السؤال في جزء الكتابة باللغتين العربية والإنجليزية.

• في كلمات مختصرة صف شعورك أثناء أدائك للاختبار ؟
 كان لي شعور جيد جداً أثناء ذلك. هذا الاختبار كان من أجلى
 تجربة. لقد تعلمت الكثير من هذا الاختبار. جميعه معي
 تمام بالوقت ولا يخلو من شيء.

• في حالة شعورك بالتوتر أثناء الاختبار في أي جزء كان هذا الشعور ؟

• هل اختلف هذا الشعور من جزء لآخر ؟
 نعم. كان شعوري جيداً في البداية، ولكن عندما
 بدأ الجزء الثاني، شعرت بالتوتر والقلق، ولا سيما في
 الجزء الأخير.

APPENDIX 14:

BERA (BRITISH EDUCATIONAL RESEARCH ASSOCIATION) GUIDELINES ON ETHICS IN CONDUCTING RESEARCH

INTRODUCTION

The British Educational Research Association adopted the following set of ethical guidelines as its Annual General Meeting on 28 August 1992. These are based on guidelines developed at the BERA seminar in March 1998 (published in *Research Intelligence*, February 1989) and the proposed ethical standards of the American Educational Research Association as published in *Educational Researcher*, December 1991. (We are grateful to the AERA Committee on Standards for permission to adapt their guidelines).

THE GUIDELINES

- 1 The British Educational Research Association believes that all educational research should be conducted within an ethic of respect for persons, respect for knowledge, respect for democratic values, and respect for quality of educational research.

Responsibility to the research profession

- 2 Educational researchers should aim to avoid fabrication, falsification, or misrepresentation of evidence, data, findings, or conclusions.
- 3 Educational researchers should aim to report their findings to all relevant stakeholders and so refrain from keeping secret or selectively communicating their findings.
- 4 Educational researchers should aim to report research conceptions, procedures, results, and analyses accurately and in sufficient detail to allow other researchers to understand and interpret them.
- 5 Educational researchers should aim to decline requests to review the work of others when strong conflicts of interest are involved or when such requests cannot be conscientiously fulfilled on time. Materials sent for review should be read in their entirety and considered carefully, with evaluative comments justified with explicit reasons.
- 6 Educational researchers should aim to conduct their professional lives in such a way that they do not jeopardize future research, the public standing of the field, or the publication of results.

Responsibility to the participants

- 7 Participants in a research study have the right to be informed about the aims, purposes and likely publication of findings involved in the research and of potential consequences for participants, and to give their informed consent before participating in research.
- 8 Care should be taken when interviewing children and students up to school leaving age; permission should be obtained from the school, and if they so suggest, the parents.
- 9 Honesty and openness should characterize the relationship between researcher, participants and institutional representatives.

- 10 Participants have the right to withdraw from a study at any time.
- 11 Reserachers have a responsibility to be mindful of cultural, religious, gendered, and other significant differences within the research population in the planning, conducting, and reporting of their research.

RESPONSIBILITY TO THE PUBLIC

- 12 Educational researchers should communicate their findings and the practical significance of their research in clear, straightforward, and appropriate language to relevant research populations, institutional representatives, and other stakeholders.
- 13 Informants and participants have a right to remain anonymous. This should be respected when no clear understanding to the contrary has been reached. Researchers are responsible for taking appropriate precautions to protect confidentiality of both participant and data. However, participants should also be made aware that in certain situations anonymity cannot be achieved.

RELATIONSHIP WITH FUNDING AGENCIES

- 14 The Data and results of a research study belong to the researcher who designed and conducted the study unless alternative contractual arrangements have been made with respect to either the data or the results or both.
- 15 Educational researchers should remain free to interpret and publish their findings without censorship or approval from individuals or organisations, including sponsors, funding agencies, participants, colleagues, supervisors, or administrators. This understanding should be conveyed to participants as part of the responsibility to secure informed consent. This does not mean however the researchers should not take every care to ensure that agreements of publication are reached.
- 16 Educational researchers should not agree to conduct research that conflicts with academic freedom, nor should they agree to undue or questionable influence by government or other funding agencies. Examples of such improper influence include endeavours to interfere with the conduct of research, the analysis of findings, or the reporting of interpretations. Researchers should report to BERA attempts by sponsors or funding agencies to use any questionable influence, so that BERA may respond publicly as an association on behalf of its members thereby protecting any individual or contract.
- 17 The aims and sponsorship of research should be made explicit by the researcher. Sponsors or funders have the right to have disclaimers included in research reports to differentiate their sponsorship from the conclusions of the research.
- 18 Educational researchers should fulfil their responsibilities to agencies funding research, which are entitled to an account of the use of their funds, and to a report of the procedures, findings and implications of the funded research.
- 19 The host institution should appoint staff in the light of its routine practices and according to its normal criteria. The funding agency may have an advisory role in this respect, but should not have control over appointments.
- 20 Sponsored research projects should have an advisory group consisting of representatives from those groups an agencies which have legitimate interest in the area of inquiry. This advisory group should facilitate access of the researcher(s) to sources of data, other specialists in the field and the wider educational community.

- 21 The funding agency should respect the right of the researcher(s) to keep his or her sources of data confidential.
- 22 In the event of a dispute between the funding agency and researcher(s) over the conduct of the research, or threatened termination of contract, the terms of the dispute and/ or grounds for termination should be made explicit by the funding agency or researcher and be open to scrutiny by the advisory group. If either party feels that grounds for termination are unreasonable then there should be recourse to arbitration by a body or individual acceptable to both parties.

PUBLICATION

- 23 Researcher(s) have a duty to report both to the funding agency and to the wider public, including educational practitioners and other interested parties. The right to publish is therefore entailed by this duty to report. Researchers conducting sponsored research should retain the right to publish the findings under their own names. The right to publish is essential to the long-term viability of any research activity, to the credibility of the researcher (and of the funding agency in seeking to use research findings) and in the interests of an open society. The methodological principle of maximising the dissemination of information to all interested parties is an integral part of research strategy aimed at testing on a continuous basis the relevance, accuracy and comprehensiveness of findings as they emerge within the process of inquiry.
- 24 The conditions under which the right to publish might be legitimately restricted are:
- general legislation (e.g. in the area of level or race relations);
 - undertakings given to participants concerning confidentiality and generally not to cause unnecessary harm to those affected by the research findings; and
 - failure to report findings in a manner consistent with the values of inquiry i.e. to report findings honestly, accurately, comprehensively, in context, and without undue sensationalisation.
- 25 Publications should indicate whether or not they are subject to reporting restrictions.
- 26 The researcher(s) should have the right, as a last resort and following discussion with the funding agency and advisory group, to publicly dissociate themselves from misleading selective accounts of the research.
- 27 Funding bodies should not be allowed to exercise restrictions on publication by default, e.g. by failing to answer requests for permission to publish, or by undue delay.
- 28 Resources need to be made available for dissemination and publication and should be built in to funding.
- 29 In the event of a dispute over publication, the researcher should seek recourse first to advisory group and secondly to an independent arbitration body or individual.

INTELLECTUAL OWNERSHIP

- 30 Authorship should be determined on the basis that all those, regardless of status, who have made a substantive and/or creative contribution to the generation of an intellectual product are entitled to be listed as authors of that product. (Examples of creative contributions are: writing first drafts or substantial portions; significant rewriting or substantive editing; contributing generative ideas or basic conceptual schema or analytic categories; collecting

data which requires significant interpretation or judgement; and interpreting data).

- 31 First authorship and order of authorship should be the consequence of relative leadership and creative contribution.

Relationship with host institution

- 32 Institutions should both develop their own codes of practice which govern ethical principles and establish appropriate standards of academic freedom, including the freedom to disseminate research findings. While such codes should be observed within all research, including non-contract research, they are particularly important in respect of contract research. Such codes should be honoured by institutions and researchers in the negotiation of contractual arrangements put forward by funding agencies, and in the carrying out of these obligations once they have been agreed.

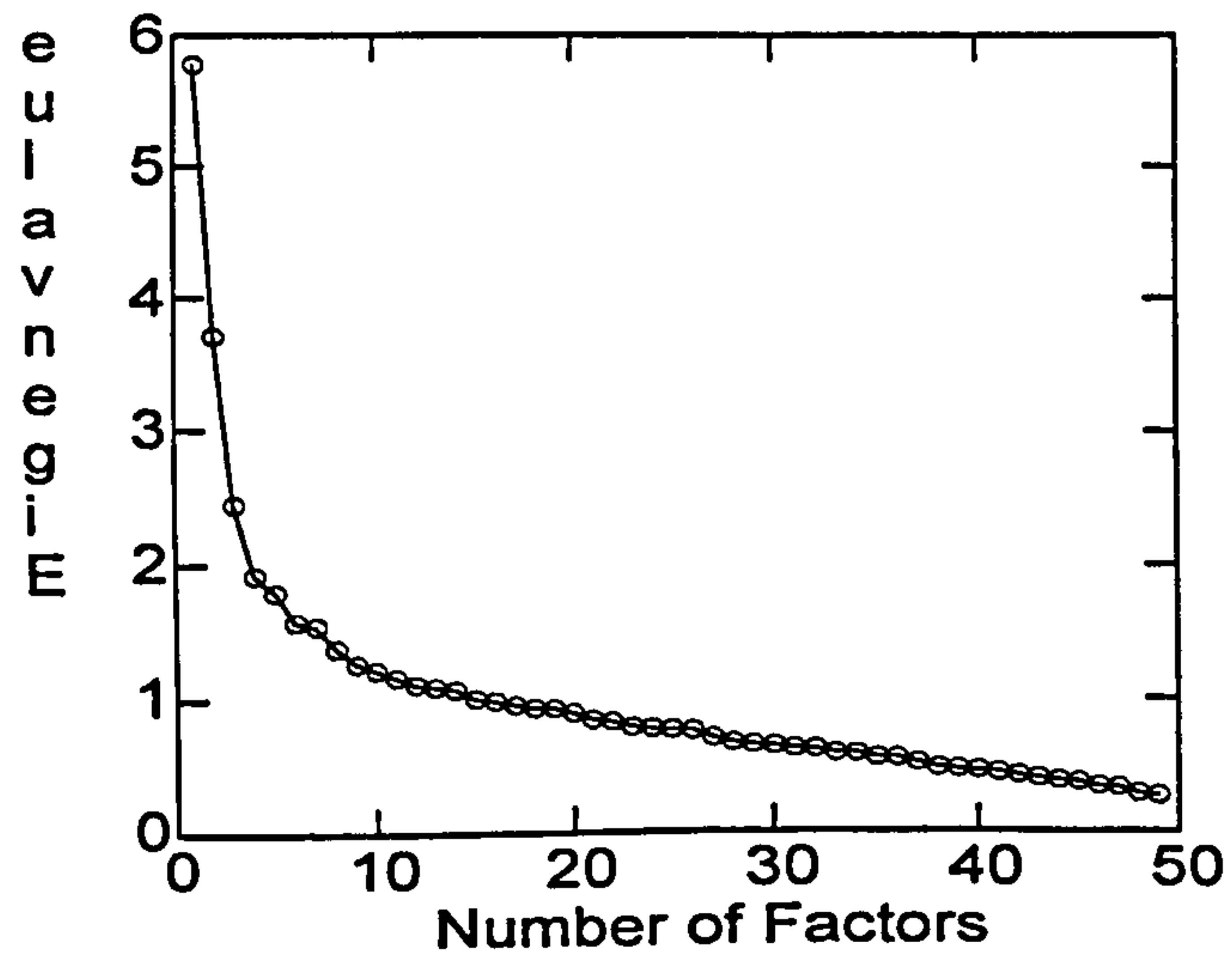
- 33 While academic staff should not engage in contract research without agreement by the institution, the latter should not be allowed to compel academic staff to engage in particular contract research.

- 34 It is assumed that contracts will in all cases be interpreted reasonably and with regard to due process. However should a legitimate disagreement arise between the funding agency and the researchers. engaged on it, then the researchers' institutions should give the researchers full and loyal support in resolving this disagreement.

APPENDIX 15:
SCREE PLOTS OF THE FACTOR ANALYSIS

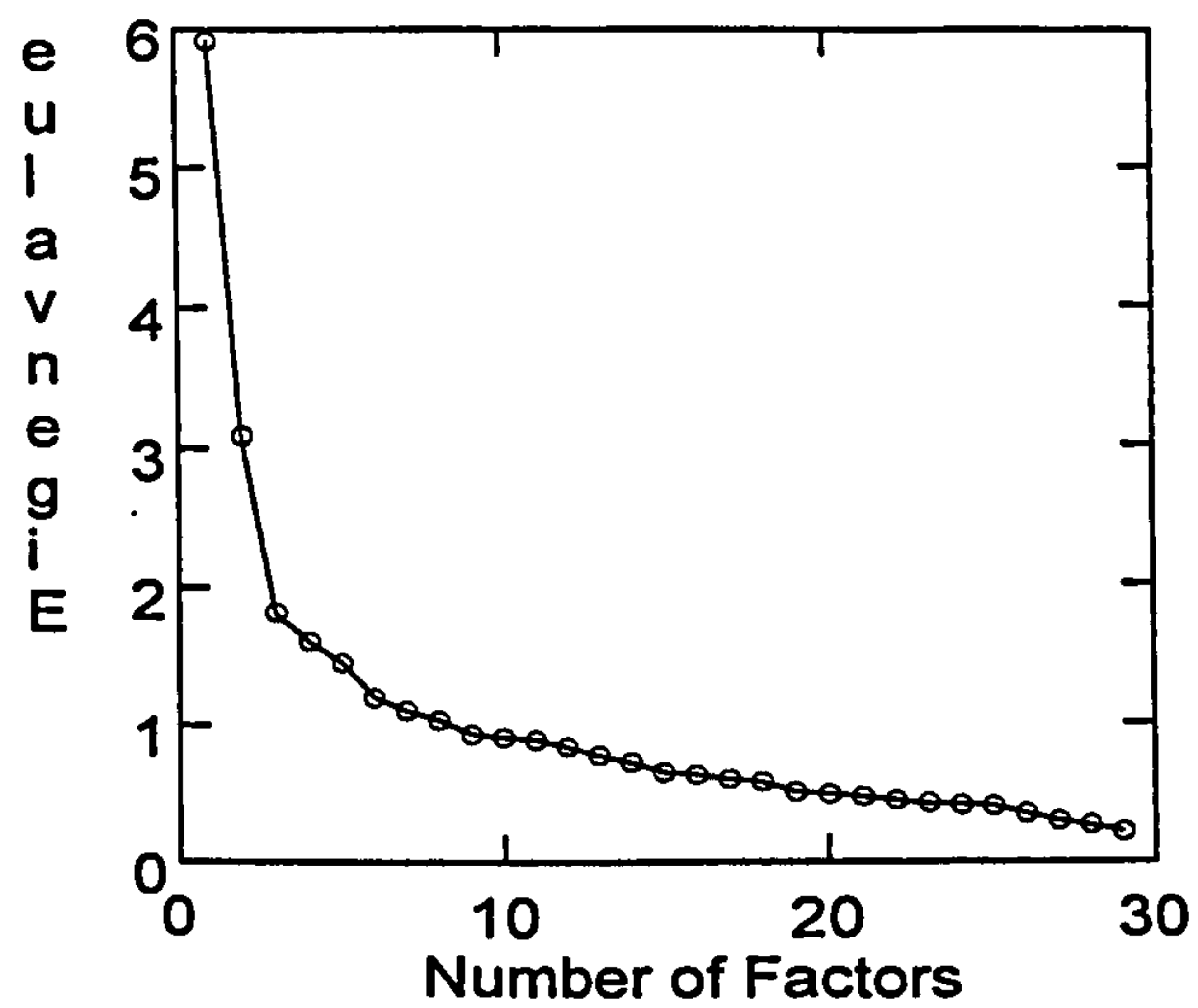
Scree Plot for the Test-Taking Strategies

Scree Plot



Scree Plot for the Emotional Regulation Processes

Scree Plot



APPENDIX 16:
DESCRIPTIONS OF THE FOUR PROFICIENCY LEVELS
AT THE CENTER FOR ADULT & CONTINUING EDUCATION
AT THE AMERICAN UNIVERSITY IN CAIRO

NOVICE

At the end of this stage, students can generally meet courtesy, introduction and identification requirements both socially and in the workplace. They are able to handle and write very familiar work-related and personal material.

ELEMENTARY

At the end of this stage, students can participate in short conversations and handle routine workplace interactions. They are able to handle material on familiar work-related and general topics and write information on simple forms and work-related documents.

INTERMEDIATE

At the end of this stage, students are able to participate in most conversations, both socially and in the workplace. They can handle factual material on general and / or academic topics and are able to meet essential workplace writing needs.

ADVANCED

At the end of this stage, students are able to use the language effectively for social and work-related conversations. They can handle materials on general and / or academic topics and are able to write most types of work-related correspondence.

**LEARNERS' PERCEPTIONS OF LANGUAGE PROFICIENCY, LANGUAGE
TEST-TAKING STRATEGIES AND EMOTIONAL REGULATION
IN A TEST-TAKING CONTEXT:
A CASE STUDY IN AN EGYPTIAN EFL CONTEXT**

DEENA BORAIE, M.A. TEFL, M.Sc.

A thesis submitted in partial fulfilment of the
requirements of the University of Wolverhampton
for the degree of Doctor of Philosophy

July 2003

This work or any part thereof has not previously been presented in any form to the University or to any other body whether for the purposes of assessment, publication or for any other purpose (unless previously indicated). Save for any express acknowledgments, references and/or bibliographies cited in the work, I confirm that the intellectual content of the work is the result of my own efforts and of no other person.

The right of Deena Boraie to be identified as author of this work is asserted in accordance with ss.77 and 78 of the Copyright, Designs and Patents Act 1988. At this date copyright is owned by the author.

Signature.....

Date.....